

GEORGES DICKER



DESCARTES

AN ANALYTICAL AND HISTORICAL INTRODUCTION

SECOND EDITION

Descartes

Descartes

An Analytical and Historical Introduction

SECOND EDITION

GEORGES DICKER

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford.

It furthers the University's objective of excellence in research,
scholarship, and education by publishing worldwide.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and certain other countries.

Published in the United States of America by
Oxford University Press
198 Madison Avenue, New York, NY 10016

© Georges Dicker 2013

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization. Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above.

You must not circulate this work in any other form
and you must impose this same condition on any acquirer.

Library of Congress Cataloging-in-Publication Data

Dicker, Georges, 1942-

Descartes: an analytical and historical introduction / Georges Dicker. -2nd ed.
p. cm.

Includes bibliographical references.

ISBN 978-0-19-538032-3 (pbk : alk. paper)—ISBN 978-0-19-970160-5 (updf)

1. Descartes, René, 1596–1650. *Meditationes de prima philosophia*. 2. First philosophy.
3. God-Proof, Ontological. 4. Methodology. 5. Knowledge, Theory of. I. Title.
B1854.D53 2013
- 194—dc23 2012030580

1 3 5 7 9 8 6 4 2

Printed in the United States of America
on acid-free paper

*To Alvina,
Beloved wife and best friend*

PREFACE TO THE SECOND EDITION

In the two decades since the first edition of this book appeared, I have continued to think about and regularly to teach Descartes's *Meditations*. I have also had the benefit of discussing my ideas about the *Meditations* and related Cartesian texts with many colleagues and students, and of reading some fine recent work on Descartes. My continued engagement with that ever-fascinating thinker has led to an expansion of knowledge, to some changes of judgment, and to a deeper understanding. In this new edition, I share these fruits of my journey.

The most significant expansion is a new chapter on the *Fourth Meditation*. In the first edition, I only summarized very briefly that *Meditation*'s main theme before moving on to the *Fifth Meditation*. Here, I offer a full-scale treatment, targeting especially the issue of how assent to a clearly and distinctly perceived truth can be, as Descartes maintains, a free act of will if, as he also maintains, such assent is unavoidable. In light of the theodicean concerns of the *Fourth Meditation*, and of the partly pedagogical aims of the book, I also provide a substantial discussion of the traditional problem of evil.

The most significant change of judgment comes in my treatment of Descartes's *cogito*. In the first edition, I argued that this most-famous of Descartes's demonstrations could in the end not avoid the dilemma of being either question-begging or invalid. Here, partly as a result of an extended correspondence with Gary Iseminger, I defend a more sympathetic assessment of the *cogito*, while still incorporating important elements of my earlier discussion such as Descartes's grounding of the *cogito* (or at least one version of it) in the theory of substance.

The main exegetical, interpretative, and evaluative themes of other parts of the book are in general the same as in the first edition. But

throughout, I have tried to pass the text through the filter of my current thinking and to make needed revisions without comprising the integrity of the original work. The result is that in virtually every section, some material has been reworked or updated, usually with a view to currency, accuracy, clarification, or completeness. For example, for currency's sake I have eliminated the account of the now-discredited "memory defense" against the accusation of the Cartesian Circle, and substituted for it a more plausible counterpart that sees the mere pastness of a clear and distinct perception as the source of its doubtfulness absent the divine guarantee of the truth of clear and distinct perceptions; for accuracy's sake, I have provided what I regard as more faithful analyses of Descartes's rationale for dismissing the insanity hypothesis in the *First Meditation* and of his notion of eminent containment in the Third and *Sixth Meditations*; for clarity's sake, I have streamlined the presentation of the problem of the Circle and sharpened the defense of the solution to it that I favor; for completeness's sake, I have enriched my discussion of the substance theory and of the alternatives to it in chapter 2, discussed the question of whether the unreconstructed cogito needs an additional, general premise, added a more text-based discussion of Kant's objection to the Ontological Argument and some discussion of Descartes's "modal" version of that argument in chapter 5, and enriched the discussion of primary and secondary qualities and of mind-body issues in chapter 6.

Another difference is that this edition no longer includes the text of the bulk of the *Meditations*. Given the availability of numerous inexpensive editions of that work, and the augmented length and expense involved in reprinting the premier English translation of it by John Cottingham, this change seemed advisable.

The passage of time has not erased my debt to the many teachers, mentors, colleagues, students, and friends, some of whom are now sadly gone, who gave me advice, support, or inspiration in writing the first edition. They included William H. Hay and Marjorie H. Stewart, to whom the first edition was dedicated, as well as Jonathan Bennett, José Bernadete, Arthur Bierman, Roland P. Blum, Roderick M. Chisholm, Fred Dretske, Richard Feldman, Robert Gemmett, Jack Glickman, Eli Hirsch, Brian O'Neil, Ingmar Persson, William L. Rowe, Marcus G. Singer, Ellen Suckiel, George J. Stack, James Syfers, James Van Cleve, Rudolph H. Weingartner, and Paul Ziff. To these I must now add people who have given me valuable comments in writing or in conversation since the appearance of the first edition, including Jean-Marie Beyssade, Krasimira

Filcheva, Richard Glauser, Gary Iseminger, Marie Jayasekera, and Peter Millican, and people whose kind remarks encouraged me to work on a new edition, including Gordon Barnes, Andrew Chignell, Gary Hatfield, James Mahon, Catherine McKeen, Keith McPartland, Peter Ohlin, Elliott Sober, and Margaret Wilson. I also thank my beloved wife, Alvina Greenberg, for her unwavering support and for preparing the line drawings that grace several pages of this book. Finally I am grateful to my caring and loyal son, Keith.

Pittsford, New York

G. D.

October 2012

PREFACE TO THE FIRST EDITION

Descartes's *Meditations* speaks to the philosophical novice as well as the sophisticate; it introduces basic issues of philosophy in way that is brief, compelling, and penetrating, and it develops them with a subtlety that remains exhilarating to us, Descartes's philosophical descendants. No wonder, then, that the *Meditations* continues to be read and analyzed at all levels of the philosophy curriculum, from the introductory course to the graduate seminar.

Like a number of other books on Descartes's philosophy, this work is essentially a commentary on his masterpiece. But unlike many of those books, it is addressed to students of the *Meditations* at virtually all levels, and to general readers interested in philosophical issues and their history.

In order to address the introductory student, I have sought not only to provide some historical background, but also and especially to elicit a number of basic issues and concepts from the *Meditations*—to milk that great text for central philosophical ideas. For example, in analyzing the *cogito* in chapter 2, I try to relate it in a systematic way to the issue of substance and identity through change and even to the problem of universals, no less than to the dualistic view of persons to which it serves as the point of entry. In further discussing that view of persons in chapter 6, I not only analyze Descartes's case for dualism but also present the problem of interaction and some historical and contemporary responses to it. Again, in discussing Descartes's views about the material world in chapter 6, I not only expound his proof of the material world as an attempt to answer the skeptical doubts generated in *Meditation I* and against the background of his arguments for God's existence but also discuss his views about the nature of material things, including the theory of primary and secondary qualities (where I draw some comparisons with

Locke). I have tried to do these things in a language and style accessible to today's college students, yet without sacrificing rigor. For the introductory student's sake, I have also tried, especially at strategic points in chapter 1, to explain briefly some matters that would be taken for granted in a book addressed solely to advanced readers, including some elementary points of logic and such things as the *a priori–a posteriori* distinction.

For students who are studying Descartes at the next higher level—typically in a survey of modern philosophy course—I have sought to cover, in a balanced way, the main themes and arguments of the *Meditations*, as well as the main criticisms that they have evoked. For example, chapter 3 presents a detailed reconstruction and a critique of Descartes's main causal argument for God's existence, as well as a critical survey of the main positions on the vexed problem of the Cartesian Circle; chapter 5 offers an extensive analysis of the *Meditation V* Ontological Argument for God's existence and the main objections that have been raised against it; and chapter 6 provides a reconstruction of Descartes's proof of the real distinction between mind and body and an analysis of his exchange with Arnauld over that argument.

To address upper-level undergraduates studying Descartes in courses on Rationalism (or on Descartes), I have tried to do a number of different things. First, I have sought to provide reconstructions of Descartes's central arguments that neither oversimplify them nor become unnecessarily technical; second, to discuss relevant episodes from Descartes's other writings, including the *Objections and Replies*, *Principles of Philosophy*, *Passions of the Soul*, *Discourse on the Method*, and a few of the letters; and finally, to draw upon, and sometimes discusses critically, recent English-language Descartes scholarship. Thus, upper-level undergraduates should find this book a useful research tool and gain from it some sense of the nature of contemporary Descartes scholarship.

I venture to hope that these latter attributes will also make this book useful to graduate students and to some of my peers. To such other readers of Descartes, I offer here my own reflections on several key Cartesian issues. For example, chapter 1 inquires whether Descartes's skeptical arguments in *Meditation I* are self-refuting; chapter 2 offers a reconstruction of the *cogito* from the substance theory and an assessment of the *cogito*'s force; chapter 3 provides a detailed examination of Descartes's various causal principles and their interrelations, and advances a critique of the view (first proposed by Anthony Kenny and later adopted by James Van Cleve and Bernard Williams) that Descartes's *Meditation III* argument

avoids circularity because its purpose is to vindicate only the general rule that all clear and distinct perceptions are true. It also proposes a solution to the problem of the Cartesian Circle that builds on the work of Alan Gewirth and Harry Frankfurt. Chapter 5 connects Kant's objection to the Ontological Argument to problems about negative existential statements, and explicates Caterus's objection in terms of the distinction between the formal and material modes of speech; and chapter 6 explores the implications of Descartes's view of matter as a purely extended substance for the individuation and identity-conditions of bodies and offers an overall assessment of Cartesian Dualism focusing on issues of logical versus causal independence of mind and body not usually discussed in treatments of the topic.

NOTE ON THE REFERENCES AND ABBREVIATIONS

All references to the works of Descartes are given in parentheses within the text. All quotations from Descartes are from René Descartes, *The Philosophical Writings of Descartes*, Volumes I and II, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press, 1984–1985) and from *The Philosophical Writings of Descartes*, Volume III: *The Correspondence*, translated by John Cottingham, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny (Cambridge: Cambridge University Press, 1991). References to Volume I are abbreviated as “CSM I,” references to Volume II as “CSM II,” and references to Volume III as “CSMK.”

Whenever possible and useful, page references have also been given to the following two one-volume selections from the same translation:

René Descartes, *Meditations on First Philosophy with Selections from the Objections and Replies* (Cambridge: Cambridge University Press, 1986), abbreviated as “M;”

René Descartes, *Selected Philosophical Writings* (Cambridge: Cambridge University Press, 1988), abbreviated as “SPW.”

References to M have not been given for the *Meditations*, but only for the *Objections and Replies*, because the page numbers for the *Meditations* in M are identical with the page numbers for the *Meditations* in CSM II.

Page references are given, as well, to the complete, original-language edition of Descartes’s works, René Descartes, *Oeuvres de Descartes*,

13 volumes, edited by C. Adam and P. Tannery (Revised edition, Paris: J. Vrin, 1964–1976). References to each volume are given by roman numeral and, where necessary, letter, for example, “AT X” for Volume X and “AT VIIIA” for Volume VIIIA.

All references to works other than Descartes’s are given in the notes. All works cited in the notes are listed in the bibliography.

CONTENTS

Preface to the Second Edition vii

Preface to the First Edition xi

Note on the References and Abbreviations xv

1. *Meditation I* and the Method of Doubt 3

- 1. DESCARTES'S GOAL 3
- 2. THE CARTESIAN DOUBT 10
- 3. IS THE CARTESIAN DOUBT SELF-REFUTING? 27
 - 3.1 The Deceptiveness of the Senses 27
 - 3.2 The Dream Argument 30
 - 3.3 The Deceiver Argument 32

2. *Meditation II*: The *Cogito* and the Self 39

- 1. DESCARTES'S "I AM THINKING, THEREFORE I EXIST" 39
- 2. THE CERTAINTY OF ONE'S OWN THOUGHTS 42
- 3. A PROBLEM FOR THE COGITO 47
- 4. THE SUBSTANCE THEORY 49
- 5. A RECONSTRUCTION OF THE COGITO BASED ON THE SUBSTANCE THEORY 54
- 6. CRITICAL DISCUSSION OF THE RECONSTRUCTED COGITO 58
 - 6.1 The Substance Theory and the Argument from Change 58
 - 6.2 The Corollary 64
 - 6.3 The Assumption That Thoughts Are Properties 68
 - 6.4 The Inference to "I Exist" 69
- 7. A DEFENSE OF THE UNRECONSTRUCTED COGITO 71

8. DOES THE UNRECONSTRUCTED COGITO REQUIRE AN ADDITIONAL PREMISE?	75
9. DESCARTES'S CONCEPTION OF THE SELF	80
10. CARTESIAN DUALISM	86
3. <i>Meditation III</i> : The Criterion of Truth and the Existence of God	91
1. DESCARTES'S CRITERION OF TRUTH	91
2. THE PROJECT OF <i>MEDITATION III</i>	98
3. FROM THE IDEA OF GOD TO GOD	100
3.1 The Nature of Ideas	101
3.2 Objective Reality and Formal Reality	105
3.3 The Core Argument	110
3.4 The Central Argument of <i>Meditation III</i> : The Subargument, the Core Argument, and the Sequel	114
4. CRITICISMS OF DESCARTES'S CENTRAL ARGUMENT IN <i>MEDITATION III</i>	129
4.1 The Subargument	130
4.1.1 The Precontainment Principle	130
4.1.2 Degrees of Reality	137
4.1.3 Justifying the Causal Maxim	139
4.2 The Problem of the Cartesian Circle	144
4.2.1 The Restriction of the Doubt to Past Clear and Distinct Perceptions Defense	147
4.2.2 The General Rule Defense	153
4.2.3 The Radical Doubt of Reason and the Creation of the Eternal Truths	164
4.2.4 The Validation of Reason	170
4.3 A Final Criticism of the Core Argument	176
4. <i>Meditation IV</i> : Error, Freedom, and Evil	181
1. THE ISSUES OF THE <i>FOURTH MEDITATION</i>	181
2. ERROR AND THE WILL	181
3. SOME POSSIBLE OBJECTIONS	184
3.1 Assenting and Deciding to Believe	184
3.2 Irresistibility and Freedom	186
4. THE COHERENCE OF CARTESIAN FREEDOM	190
5. DESCARTES'S TROUBLING LETTER TO MESLAND	197

6. ERROR AND EVIL	205
6.1 The Problem of Evil	205
6.2 Cartesian Theodicy	214
6.3 Some Critical Reflections	217
5. <i>Meditation V</i> : The Ontological Argument for the Existence of God	221
1. DESCARTES'S ONTOLOGICAL ARGUMENT	221
2. CRITIQUE OF THE ONTOLOGICAL ARGUMENT	228
2.1 Gaunilo's Objection	229
2.2 Kant's Objection	231
2.3 Further Consideration of Kant's Objection	237
2.4 Caterus's Objection	242
3. SOME IMPLICATIONS FOR DESCARTES'S SYSTEM	254
6 : <i>Meditation VI</i> : Dualism and the Material World	257
1. THE SCOPE OF MEDITATION VI	257
2. DESCARTES'S PROOF OF THE REAL DISTINCTION BETWEEN MIND AND BODY	258
3. DESCARTES'S PROOF OF THE MATERIAL WORLD	274
4. DESCARTES ON THE NATURE OF THE MATERIAL WORLD	282
4.1 Primary and Secondary Qualities	283
4.2 Matter, Space, and Solidity	292
4.3 Bodies as Substances versus Bodies as Modes of Substance	295
5. DUALISM AND THE PROBLEM OF INTERACTION	301
6. AN ASSESSMENT OF CARTESIAN DUALISM	313
<i>Bibliography</i>	329
<i>Index</i>	335

Descartes

Meditation I and the Method of Doubt

1. Descartes's Goal

If one had to describe in a single word what Descartes does in his *First Meditation*, that word would have to be the verb “doubt.” Throughout *Meditation I*, Descartes doubts, or calls into question, his previous beliefs. From a logical point of view, however, Descartes’s famous and dramatic decision to doubt all his previous beliefs is not his point of departure. Rather, his logical point of departure is a statement of the purpose for which he will doubt them. In the very first sentence of *Meditation I*, Descartes declares that he must question his beliefs “if I [he] wanted to establish anything at all in the sciences that was stable and likely to last,” and near the end he repeats that he must withhold assent from his previous beliefs “if I [he] want[s] to discover any certainty” (CSM II 15, SPW 79, AT VII 22). The two phrases in quotation marks are crucial, because in them Descartes states his purpose: to discover what, if anything, is really certain. It is for the sake of this goal that Descartes resolves to doubt his previous beliefs. An analogy of Descartes’s own nicely illustrates his basic strategy. Suppose that you had a basket full of apples, that you feared some of them might be rotten, and that you wanted to find the good ones. How might you proceed? Well, the easiest way would be to turn all the apples out of the basket, inspect them, and put back into the basket only the unspoiled ones (CSM II 324, SPW 123, M 63, AT VII 481). Likewise, by trying to doubt all of his beliefs, Descartes hopes to find some beliefs that he cannot doubt, that is, that are genuinely certain. Indeed, this “quest for certainty” (to borrow a phrase from the American philosopher John Dewey) is the engine that drives Descartes’s *Meditations* as a whole.

Descartes’s goal does not come as the conclusion of an argument, so it would be futile to look for some line of reasoning leading up to it. As one

contemporary commentator on Descartes points out, “it is proper for the course of an inquiry to be guided by its goal and, so far as logic is concerned, the goal of the inquiry may be postulated as a matter of free choice. . . . Descartes states his purpose in the *First Meditations*’s opening sentence.”¹ Nevertheless, before proceeding with *Meditation I*, we should ask why this goal seemed important to Descartes, and why his attempt to attain it remains of interest today.

One reason is simply that the question “What, if anything, is really certain?” is an intrinsically interesting one. We live in a world where there are diverse opinions, views, and theories about many matters. Often, these views are held with great confidence; yet what passes for knowledge is continually changing, as new scientific discoveries are made and new theories devised. Experts disagree on many important matters. Have you ever asked yourself, then, whether anything is really certain? Or are humans bound to live in a sea of uncertainty, taking their guidance from the prevailing “expert” opinion of the day—an opinion that may change by tomorrow? Descartes, for one, was profoundly dissatisfied with such a prospect. He was convinced that by a careful application of thought, genuine certainty could be found, even about matters of the greatest moment. Does the thought that perhaps nothing is certain leave you dissatisfied? Can you sympathize with Descartes’s hope that careful thinking can lead to genuine certainty about important matters? If so, then you can appreciate the most basic reason for asking what, if anything, is really certain: the intrinsic interest of the question.

Another reason pertains to the nature of knowledge. There seems to be a close connection between knowing something and being certain of it. Suppose I were to tell you, “I know that Paris is the capital of France, but I am not completely certain that it is. I admit that it’s a little doubtful in my mind whether Paris really is the capital of France.” You may well feel that there would be something very wrong (and not merely an ignorance of geography) with those statements. If I really know that Paris is the capital of France, you may want to insist, then I am also (completely) certain that it is. Now Descartes, and many other philosophers both before and after him, have maintained that knowledge and certainty indeed go hand in hand. This provides a further reason for asking what, if anything, is genuinely certain.

¹ Harry Frankfurt, *Demons, Dreamers, and Madmen*, p. 27. The first half of this book is a very helpful and readable analysis of *Meditation I*.

There are also reasons of a historical kind (though even these have modern counterparts). Let us consider two that exercised a powerful influence on Descartes: the rise of science in the seventeenth century, and the revival of philosophical skepticism.

In the seventeenth century, there occurred a series of scientific discoveries that challenged and eventually destroyed the medieval conception of the universe—a conception that had endured for nearly two thousand years. According to this conception, which was rooted in the physics of Aristotle (384–322 B.C.) and the astronomy of Ptolemy (100–170 A.D.), the universe is a sphere with the earth at its center. The moon, planets, sun, and stars all revolve around the earth in fixed, circular orbits. The universe contains two regions, the sublunar and the supralunar. The sublunar region is the interval between the moon and the earth; the supralunar region comprises everything from the moon to the outermost circumference of the universe. The difference between the two regions pertains to the kinds of *changes* that occur in each of them. In the sublunar region, there are several different sorts of changes: coming into being and passing out of being (what Aristotle called “generation and corruption”), changes in things’ qualities, changes of position (locomotion), changes in the number of things that exist. But in the supralunar region, there is only one kind of change: perfect, circular motion of the heavenly bodies around the earth (and in “epicycles” around certain points in their own orbits around earth). Furthermore, while all locomotion in the supralunar region is circular, locomotion in the sublunar region is always rectilinear motion toward what Aristotle termed a thing’s “natural place.” Heavy things like earth move downward toward their natural places; light ones like fire move upward toward theirs. The upshot is that the two regions have completely different principles of motion—completely different physics.

An additional but related element in the medieval cosmos is that of hierarchy. There is a genuine “best–better–worse” scale built into the universe. This hierarchy is no mere subjective human value judgment. Rather, it is built into the very fabric of things; for it stems from the types of changes found in the two regions. Since the sublunar region contains such changes as decline, death and decay (“corruption”), it is not as admirable as the supralunar region, where nothing ever dies or passes out of being, but rather everything exhibits only perfect, circular motion. Thus the supralunar region is better than, or superior to, the sublunar.

A final, crucial element in the medieval conception of the universe is teleology. “Teleology” comes from the Greek word *telos*, which means

“purpose.” According to the medieval conception, everything in the universe serves a purpose. Thus, when Galileo claimed to have seen the moons of Jupiter through his newly discovered telescope, it was argued against him—as a serious scientific argument—that Jupiter could not have any moons, since they would serve no purpose. The purpose of each thing throughout all of nature is, roughly, to be as like God as it is possible for that kind of thing to be. Thus, the perfect, circular motion of the stars approximates the immutability (unchangeability) of God much more closely than the more varied and chaotic changes in the sublunar region. But even in the sublunar region, where individual creatures are born, grow, decay, and die, some resemblance to God’s immutability is preserved in the fixity of species (which was not to be seriously challenged until Darwin’s *The Origin of Species* in the nineteenth century).

This conception of the universe as an earth-centered sphere exhibiting built-in hierarchy and teleology endured for almost two thousand years. It had the authority of Aristotle (who was revered as “the Philosopher” throughout the Middle Ages), of the Church, and of the universities behind it. But the scientific discoveries of the seventeenth century—first in astronomy and then in physics—challenged its details and eventually ruined its ensemble. Let us glance at a few examples. The astronomer Tycho Brahe showed that comets, which were assigned to the supralunar region, are not indestructible (they burn up), thereby providing an example of “corruption” where no such change was supposed to occur. He also discovered sunspots—another example in the supralunar region of changes (qualitative ones) that were supposed to occur only in the sublunar region. Kepler, building on observations made by Brahe, showed that the planets do not move with an even circular motion, but an uneven elliptical one. Galileo’s experiments refuted Aristotle’s view that all objects in the sublunar region move in a straight line toward their “natural places.” These and other discoveries undermined the hierarchical conception of the universe, since this conception was rooted in the difference between the types of changes in the “inferior” sublunar and the “superior” supralunar regions, that is, in the idea that the physics of the two regions are entirely different. Finally, as Copernicus’s theory—that the earth is not static and at the center of the universe, but instead revolves around the sun—gained ever greater acceptance, it ruined the entire medieval cosmos.

The teleological conception of nature did not escape the onslaught of the new science either. Consider just the implications of Newton’s First

Law, which had been anticipated by Galileo. According to that law, a body remains in its state of motion or rest unless some force acts upon it. The implication is that in order to explain why a body accelerates or decelerates, no reference to purpose is required or relevant. Putting the matter crudely, all that is required is a reference to the push, pull, or gravitational attraction of some other body. By generalizing from this example, one can gain some appreciation of the transformation that resulted once teleology was expelled from nature. Now, all physical changes were to be explained in terms of mathematically formulable laws that made no reference to purpose, rather than in terms of things' striving to realize a purpose inherent in nature. Formerly, the universe could be conceived on the analogy of a giant living organism striving toward a goal. Now, it would be conceived on the analogy of a huge machine operating in accordance with purely mechanical principles.

While the conception of the universe that had reigned virtually uncontested for centuries was being undermined by the new science, the old certainties were being eroded from another direction as well. In the late sixteenth century, there occurred a revival of philosophical skepticism led by the French essayist Michel de Montaigne (1533–1592). Skepticism has been an important tradition in philosophy since antiquity, and continues to have adherents to this day. As a philosophical position, skepticism calls into question the possibility of knowledge. Skeptics typically use certain arguments intended to show that our cognitive faculties (our senses, reason, and memory) are not adequate to enable us to distinguish between truth and falsehood, and so not adequate to enable us to obtain knowledge. Montaigne revived these arguments, which date back at least to the Greek skeptic Pyrrho (360–275 B.C.).

Most of Montaigne's arguments, like those of Pyrrho and other early skeptics, were directed against the senses. Since we shall have occasion to examine such arguments with care later, we shall not discuss them in detail now. But their general tenor can be gleaned from a few passages from Montaigne's *Essays*:

We no longer know what things are in truth; for nothing comes to us except falsified and altered by our senses. When the compass, the square, and the ruler are off, all the proportions drawn from them, all the building erected by their measure, are also necessarily imperfect and defective. The uncertainty of our senses makes everything they produce uncertain. . . . Furthermore, who

shall be fit to judge these differences? . . . If he is old, he cannot judge the perceptions of old age, being himself a party in the dispute; if he is young, likewise; healthy, likewise; likewise sick, asleep, or awake. We would need someone exempt from all these qualities . . . and by that score we would need a judge that never was. . . . To judge the appearances that we receive of objects, we would need a judicatory instrument; to verify this instrument, we need a demonstration; to verify the demonstration, an instrument: there we are in a circle. . . . Now if anyone should want to judge by appearances anyway, to judge by all appearances is impossible; for they clash with one another by their contradictions and discrepancies, as we see by experience. Shall some selected appearances rule the others? We shall have to verify this selection by another selection, the second by a third; and thus it will never be finished.²

Descartes was very familiar with these and other skeptical arguments. They provided an additional incentive for him to inquire what, if anything, is really certain. As we shall see, he attempted to refute skeptical arguments once and for all, by first carrying them much further than anyone had previously done, and then showing that even his radicalized versions of the arguments could be answered.

To understand Descartes's quest for certainty, one further factor needs to be mentioned: Descartes was a mathematical genius. He discovered analytical geometry and invented the Cartesian coordinates (which are named after Cartesius, the Latinized version of "Descartes"). While Descartes was still in secondary school (he attended a Jesuit college named La Flèche), he came to feel that most of what he was being taught was not genuine knowledge. At the same time, he was impressed and delighted with the clarity and certainty that he found in mathematics. Accordingly, he conceived the idea that all genuine knowledge ought to be as clear and certain as mathematical knowledge. This became, for Descartes, the fundamental requirement for knowledge: it must be as certain as geometry or algebra. Only so would it be immune to the skeptics' attacks.

In order to fulfill this requirement, Descartes devised a method, which he elaborated at length in an early work entitled *Rules for the Direction of the Mind* and in his famous *Discourse on the Method*. We need not go into

² Michel de Montaigne, *Essays and Selected Writings*, pp. 245–47.

the details of this method, but shall only state its most basic rule. This rule directs us to accept no propositions as true except (a) those which are so obvious and clear that they cannot be doubted so long as one is thinking of them attentively and (b) those which logically follow from propositions of kind (a). In other words, certainty is to be attained by making sure, as in mathematical proof, that knowledge has the pattern of a deductively valid argument starting from self-evident, unshakeable premises. As the following passage from the *Discourse on the Method* shows, Descartes had high hopes for this method:

Those long chains composed of very simple and easy reasonings, which geometers customarily use to arrive at their most difficult demonstrations, had given me occasion to suppose that all the things which can fall under human knowledge are interconnected in the same way. And I thought that, provided we refrain from accepting anything as true which is not, and always keep to the order required for deducing one thing from another, there can be nothing too remote to be reached in the end or too well hidden to be discovered. (CSM I 120, SPW 29, AT VI 19)

The purpose of the doubt that Descartes adopts in *Meditation I* can now be clarified. Its purpose is to determine what propositions, if any, cannot be doubted. For if Descartes can find such indubitable propositions, then he will have the propositions of kind (a) mentioned above, from which he can then hope to deduce propositions of kind (b). By his doubt, then, Descartes does not mean to reject permanently all of his former beliefs. Some of them may well be true. But if they are, then Descartes wants to rediscover them, in the sense of showing that they follow logically from basic, indubitable propositions. The main purpose of the doubt is to find these indubitable propositions, so that Descartes can use them as “foundations” upon which to rebuild his knowledge. The doubt is a way of rethinking everything from the beginning, so as to achieve the certainty that Descartes is seeking.

In this section, we have sought to understand why Descartes embarks on his famous and seminal quest for certainty. We have seen that the question “What, if anything, is certain?” is an intrinsically interesting one and that it seems to be closely related to the question “What, if anything, do we really know?” We have also come to an appreciation of why Descartes, a mathematical genius living at a time when old certainties

were being shaken by the new science and skeptical philosophers were renewing their corrosive attacks on the very possibility of knowledge, should have adopted the quest for certainty as his basic goal. Finally, we have briefly sketched the method that his mathematical pursuits inspired him to devise for attaining that goal. At this point, then, we turn to an examination of the text of *Meditation I*, where Descartes puts this method to work.

2. The Cartesian Doubt

In this section, we shall follow closely the movement of Descartes's thought in *Meditation I*, by constructing a step-by-step summary of the *Meditation*. As we have already seen, Descartes begins with a statement of his purpose. We may accordingly enter the following statement as the first step of our summary:

1. I want to discover what, if anything, is absolutely certain.

Shortly after stating this goal, Descartes declares that "reason now leads me to think that I must hold back my assent from opinions which are not completely certain and indubitable just as carefully as I do from those which are patently false" (CSM II 12, SPW 17, AT VII 18). To say that reason "leads me to think that . . ." is to imply that one is reasoning from some premise. Then from what premise is Descartes here reasoning? The answer is clear: from his statement of purpose. For given that one's purpose is to find absolute certainty, one has an excellent reason not to accept things that are uncertain! Descartes is here reasoning directly from his goal to what he must do in order to attain it. Thus, we may enter the following statement as the second step in our summary:

2. For this purpose, I must withhold belief from things that are not entirely certain and indubitable just as carefully as from those which are obviously false.

Descartes here declares that he will be just as wary of what is even slightly doubtful as of what is obviously false. This attitude can be clarified by contrasting it with the attitude of a person concerned with practical matters—say with making money. In order to succeed, such a person

must adopt an attitude which is precisely *not* the one Descartes here adopts. He cannot be just as suspicious of what is only somewhat doubtful as of what is obviously false. Instead, he must be willing to take risks, to act on the probability that his beliefs are correct. Otherwise, he will remain paralyzed, so to speak, and never achieve his end. Suppose, however, that your only purpose is to discover what is certain. Suppose that at least for the moment, you are not concerned with any practical ends, but only with attaining absolute certainty. Then Descartes's policy of withholding belief even from matters that are only slightly doubtful is perfectly reasonable.

But what exactly does it mean to "withhold" belief (or assent)? Well, it means the same thing as does the more commonly used phrase, "to suspend judgment." To clarify this concept, notice that although there are only *two* possibilities regarding the truth of a statement (the statement is either true or false), there are *three* different postures regarding belief of a statement. These three belief-postures, or "doxastic attitudes" (to borrow a term from a contemporary American philosopher Keith Lehrer), are as follows: (1) one can believe the statement—accept it as true—(2); one can disbelieve the statement—reject it as false; (3) or one can "withhold" the statement—*neither* believe *nor* disbelieve it. To give an example, we can compare the doxastic attitudes of a theist, an atheist, and an agnostic toward the statement "God exists." A theist is someone who believes "God exists." An atheist is someone who disbelieves "God exists." An agnostic is someone who withholds (belief in) "God exists" (neither believes nor disbelieves it). As this example illustrates, there is a big difference between disbelieving and withholding. Disbelieving a statement, "*p*," is the same thing as believing its denial or negation, "*not-p*": the atheist, who disbelieves "God exists," thereby believes "God does not exist." On the other hand, withholding (belief in) a statement commits one neither to the statement itself nor to its negation: the agnostic believes neither "God exists" nor "God does not exist." Withholding, then, is a neutral, noncommittal attitude, by which one avoids committing oneself to the truth of either a statement or its denial.

We can now see more clearly why Descartes withholds (belief in) statements that are uncertain. He decides to adopt a policy that will never allow him to accept any statement that is uncertain. So which of the three "doxastic attitudes" must he take toward statements that are uncertain? Well, obviously, he must not believe them. Should he then disbelieve them? No. For then he would believe their negations, which would go

against his policy; for often the negation of an uncertain statement is itself uncertain (often both p and $\neg p$ are uncertain). For example, “there will be a major earthquake in California next year” is uncertain, but so is “there will not be a major earthquake in California next year.” Therefore, the only policy toward uncertain statements that can never lead to accepting an uncertain statement is that of withholding belief.

One can even give a plausible argument showing that when a statement is uncertain, its negation is *never* certain. This argument (which, incidentally, is not given by Descartes himself, and which you can skip over without losing track of this exposition of the *First Meditation*) depends on the following classification of statements. Statements may be classified as either certain or not certain, and those that are not certain may be further subdivided into statements that are uncertain and statements that have no credibility whatsoever. These last two classes do not overlap, for to call a statement “uncertain” is to concede that it has *some* degree of credibility. Only as a joke could one say that a statement that has no credibility whatsoever, such as $1 + 1 = 3$, is “uncertain”; for this statement is not (merely) uncertain, it is absurd or obviously false. So, although it belongs in the class of statements that are not certain (since otherwise it would have to be classed as certain), it does not belong to the subclass of that class consisting of uncertain statements. Rather, it belongs in the subclass (of the class of statements that are not certain) consisting of statements that have no credibility. If this classification of statements is correct, then one can construct a simple argument showing that if a statement p is uncertain, then its negation, $\neg p$, is not certain. The first premise is:

(1) If $\neg p$ is certain, then p has no credibility.

The only other premise is:

(2) If p has no credibility, then p is not uncertain.

To see the plausibility of these two premises, let p stand for some obvious falsehood, such as $1 + 1 = 3$. Then (1) seems indisputably true. Further, in light of the distinction between statements that are altogether lacking in credibility and statements that are merely uncertain, (2) also seems to be true. But from (1) and (2), it follows that:

(3) If $\neg p$ is certain, then p is not uncertain.

The step from (1) and (2) to (3) is valid, because it has the following obviously valid form (called *Hypothetical Syllogism*):

If P, then Q

If Q, then R

. \therefore If P, then R

To see that the step has this form, substitute “not-p is certain” for P, “p has no credibility” for Q, and “p is not uncertain” for R. Finally, it follows from (3) alone that:

(4) If p is uncertain, then not-p is not certain.

The step from (3) to (4) is valid, because it has the following form:

If P, then not-Q

. \therefore If Q, then not-P

To see that the step from (3) to (4) has this form, substitute the statement “not-p is certain” for P and the statement “p is uncertain” for Q. To see that the form itself is valid, consider a simple example:

(3a) If this is a triangle, then this does not have four sides.

. \therefore (4a) If this has four sides, then this is not a triangle.

Another way to see why the step from (3) to (4) is valid is to apply to it two rules of logic. One is *Contraposition*, which says that

If P, then Q.

. \therefore If not-Q, then not-P.

is a valid form of argument. The other is *Double Negation*, which says that not not-P is equivalent to P (e.g., “today is not not Wednesday” is equivalent to “today is Wednesday”). Substitute “not-p is certain” for P and “p is not uncertain” for Q. Then applying *Contraposition* to (3) yields “if p is not not uncertain, then not-p is not certain”; and applying *Double Negation* to

this last statement yields (4). The above argument, then, appears to be sound. If that is right, then the argument provides an additional reason why Descartes must adopt a policy of withholding rather than disbelieving uncertain statements: their negations are *never* certain.

But how is Descartes's policy to be implemented? Obviously, he cannot examine all of his beliefs individually: that would be "an endless task." Accordingly, he proposes to examine the *basic principles* on which his beliefs rest. For if these principles are uncertain, then so are any beliefs resting on them. We may enter the following statement, then, as the third step in our summary:

3. To do this, I need not examine all my beliefs individually, but only the basic principles on which they rest.

In order to understand what comes next, it is important to realize that in his *Meditations* Descartes means to be speaking not only for himself but for anyone who is seeking to determine the certainty of his or her beliefs. Despite the fact that Descartes's entire *Meditations* is written in the first person singular, the work is certainly not intended as a report of one man's idiosyncratic musings. Rather, Descartes means to speak for all of us. He is convinced that anyone who embarks upon the quest for certainty methodically and without becoming confused will travel the same route as he himself does in his *Meditations*. Indeed, some recent commentators seek to convey the impersonality of the "I" in the *Meditations* by always referring to Descartes in the third person, as "the meditator," and even to convey its gender neutrality by referring back to the meditator as "she."³ We shall not follow this practice, but simply note here that the "I" of the *Meditations* is as much an invitation to the reader to put himself or herself in Descartes's place, as it is a way for Descartes to report his own progress.

Accordingly, suppose that we each ask ourselves, on what basic principles are *my* beliefs based? Moreover, in asking this question, let us not target our more theoretical or esoteric beliefs, like beliefs about subatomic particles, or outer space, or the distant past. Let us target, instead, those beliefs that we hold most confidently and unhesitatingly. What sorts of beliefs are these? Well, they are beliefs about our present, immediate physical surroundings—such as your belief that there

³ See, for example, Gary Hatfield's excellent *Descartes and the Meditations*.

is a book on your desk or that someone is speaking to you. What are such beliefs based on? The answer is obvious: they are based on your present perceptions of sense—on what you now see, hear, feel, and so on. If someone asked why you believe there is a book on your desk, the answer would be “Because I see it” or “Because I feel it”; if someone asked why you believe someone is speaking to you, the answer would be “Because I hear him” or “Because I see him.” The beliefs that we accept as most obvious and certain, then, are based on our present perceptions of sense. The “basic principle” on which they rest is that the senses provide us with highly reliable information about our physical surroundings. This is exactly what Descartes, speaking not idiosyncratically but for all of us, goes on to say: “Whatever I have up till now accepted as most true I have acquired either from the senses or through the senses” (CSM II 12, SPW 76, AT VII 18). Let us enter this step into our summary, as follows:

4. Until now, everything I've accepted as most obvious and certain has been based on the senses.

It is one thing, however, to say that something is accepted as certain, and another to say that it really is certain. Can we assert, then, that beliefs acquired by using our senses really are certain? No, answers Descartes—at least not without qualification; for sometimes the senses are deceptive. Descartes gives no examples of this deceptiveness in *Meditation I*, perhaps because the skeptics had made such examples so familiar to Descartes's contemporaries that he judged it unnecessary to do so. When he reviews this point in *Meditation VI*, however, he does give an example, which we may usefully mention now:

Sometimes towers which had looked round from a distance appeared square from close up; and enormous statues standing on their pediments did not seem large when observed from the ground. In these and countless other such cases, I found that the judgements of the external senses were mistaken. (CSM II 53, SPW 113, AT VII 76)

The occasional deceptiveness of the senses, then, provides a reason not to trust them uncritically. As Descartes puts it: “But from time to time I have found that the senses deceive, and it is prudent never to trust

completely those who have deceived us even once" (CSM II 12, SPW 76, AT VII 18). Let us enter this point as the next step of our summary:

5. But the senses have sometimes deceived me; I'd better not trust them completely.

With this step, Descartes introduces a theme that occupies him throughout the rest of *Meditation I*: a critique of the senses. By this critique, Descartes wants to determine to what extent, if any, the senses provide certainty. He will argue, as we shall see, that they provide no certainty. Thus while the primary purpose of the Cartesian doubt is to find certainty, it also has a secondary purpose: to show that certainty is not provided by the senses—thereby preparing the way for Descartes's own views about how certainty can be attained, which he will present in the subsequent *Meditations*.

In his critique of the senses, however, Descartes wishes to be scrupulously fair. He knows that despite what skeptics may say, we are all very strongly inclined to believe that our senses provide us with highly reliable—indeed certain—information. As the saying goes, "Seeing is believing." To challenge this deep-seated conviction in a way that will lead us to reconsider it seriously, Descartes takes into account what can be said on behalf of the senses. He tries, so to speak, to give the senses as much credit as they are due. This desire to anticipate whatever can be reasonably said in defense of the senses lies behind his next point:

Yet although the senses occasionally deceive us with respect to objects which are very small or in the distance, there are many other beliefs about which doubt is quite impossible, even though they are derived from the senses—for example, that I am here, sitting by the fire, wearing a winter dressing-gown, holding this piece of paper in my hands, and so on. Again, how could it be denied that these hands or this whole body are mine? (CSM II 12–13, SPW 76–77, AT VII 18)

Descartes's point here is one that naturally occurs to anyone thoughtfully trying to weigh the impact of the fact that the senses are sometimes deceptive. This is that while the senses can indeed lead us astray in cases where the object perceived is very small or far away (as in the case of seeing a tower in the distance), it does not follow that the senses are

unreliable when the object is sizable, close by, or the like (as in the case of seeing a piece of paper in my own hands). To generalize the point, Descartes is saying that just because objects are sometimes misperceived because the *conditions of observation* are poor, it does not follow that the senses are unreliable even when the conditions of observation are good. Let us enter this important point into our summary:

6. Perceptions occurring in poor conditions of observation (e.g., when the object is minute or distant) are suspect, but can't I trust perceptions occurring in favorable conditions (e.g., of a piece of paper in my own hands as I sit by the fire)?

The position reached in step 6 would seem to be simply intelligent common sense. The general principle it embodies—that the senses may be trusted so long as the conditions of observation are good—seems eminently reasonable. Yet, doubts might be raised. For example, one might ask whether it is really enough that the conditions of observation *be* favorable: Is it not also required that the perceiver *know* that they are favorable? But how is one to know this except by other perceptions that are also potentially deceptive? And if we require that these latter perceptions also be known to occur in favorable conditions, then we shall have to appeal to still further perceptions for knowledge of those conditions, and so on without end—thereby launching ourselves into what philosophers call a “vicious infinite regress.”⁴

This line of thought, however, is not one that Descartes pursues. First, he briefly mentions, but quickly dismisses, the possibility that he might be insane, like “madmen whose brains are so damaged that . . . they firmly maintain that they are kings when they are paupers, or say they are dressed in purple when they are naked, or that their heads are made of earthenware, or that they are pumpkins, or made of glass” (CSM II 13, AT VII 19). Instead of exploring that alarming possibility, he invokes one of his most famous arguments: the *Dream Argument*. But before looking at that argument, let us pause to ask: Why does Descartes so easily dismiss the insanity hypothesis? The reason, we suggest, stems from the nature of his project. He is not trying to determine whether a person who is demented

⁴ This line of argument is developed in my “Is There A Problem About Perception and Knowledge?” 165–76, and in my *Perceptual Knowledge: An Analytical and Historical Study*, chap. 1.

or insane can discover any certainty, or much less, as he also puts it (revealing the fuller scope of his ambitions), whether such a person can “establish anything in the sciences that [is] stable and likely to last” (CSM II 12, SPW 17, AT VII 17). Nor, despite the fact that Descartes was without doubt one of the greatest geniuses who ever lived, is he asking whether a person endowed with a superior mind can discover any certainty or establish anything stable and lasting in the sciences. Rather, he is asking whether a rational mind, endowed with normal human intelligence, can, by careful and methodical thinking, achieve these ends. So, the hypothesis that he might be insane is simply irrelevant to his inquiry. Of course, someone may say that he *ought* to have considered it more seriously, but the question would then be: *why* ought he? Nothing in his project commits him to doing so, since his purpose is to determine whether a rational mind can attain certainty.⁵ Furthermore, the question of whether a demented person could attain certainty, or establish lasting results that would remain fixed and stable in her mind, seems almost to answer itself.

⁵ In this regard, consider the very beginning of Descartes’s *Discourse on the Method*:

Good sense is the best distributed thing in the world: for everyone thinks himself so well endowed with it that even those who are hardest to please in everything else do not usually desire more good of it sense than they possess. In this it is unlikely that everyone is mistaken. It indicates rather that the power of judging well and of distinguishing the true from the false—which is really what we properly call ‘good sense’ or ‘reason’—is naturally equal in all men, and consequently that the diversity of our opinions does not arise because some of us are more reasonable than others but solely because we direct our thoughts along different paths and do not attend to the same things. For it is not enough to have a good mind; the main thing is to apply it well. The greatest souls are capable of the greatest vices as well as the greatest virtues; and those who proceed but very slowly can make much greater progress, if they always follow the right path, than those who hurry and stray from it.

For my part, I have never presumed my mind to be in any way more perfect than that of the ordinary man; indeed, I have often wished to have as quick a wit, or as sharp an imagination, or as ample and prompt a memory as some others. And apart from these, I know of no other qualities which serve to perfect the mind; for as regards reason or sense, since it is the only thing that makes us men and distinguishes us from the beasts, I am inclined to think that it exists whole and complete in each of us. (CSM I 111, SPW 20–21, AT VI 2)

The first sentence is not a sarcastic remark. Descartes thinks that all sane human beings possess the power to reason well and to judge correctly (what he here calls “good sense”), and that part of possessing that power is recognizing that one possesses it. The passage as a whole reflects very well the standpoint from which Descartes wrote the *Meditations*.

On the other hand, Descartes cannot just dismiss the possibility that he might be dreaming, since it is a plain fact that even ordinary, rational humans have dreams. Thus he exclaims,

As if I were not a man who sleeps at night, and regularly has all the same experiences while asleep as madmen do when awake—indeed sometimes even more improbable ones. (CSM II 13, SPW 77, AT VII 19)

The fact of dreaming, then, provides a reason to doubt even perceptions occurring under the best conditions. This is that such perceptions can be exactly duplicated in a vivid dream. Of course, not all dreams are that life-like: some dreams have a dream-like quality. But all that is needed to provide some reason to doubt even our “best” perceptions is that *some* dreams be so realistic, so apparently authentic, as to be indistinguishable (during the time of their occurrence) from waking experience. Now it seems difficult to deny that this is the case. As Descartes puts it:

As I think about this more carefully, I see plainly that there are never any sure signs by means of which being awake can be distinguished from being asleep. The result is that I begin to feel dazed, and this very feeling only reinforces the notion that I may be asleep. (CSM II 13, SPW 77, AT VII 19)

Let us accordingly enter, as the next step in our summary, the following response to step 6:

7. No, for even perceptions occurring in ideal conditions of observation may be indistinguishable from vivid dreams.

In step 7, Descartes has reached a skeptical position with regard to the senses. This can be appreciated by comparing (7) with (5). Step 5 only implied that the senses should not be trusted completely or uncritically, since some perceptions—those that occur under adverse conditions—are deceptive. This is hardly a skeptical result; it is, as already noted, just common sense. But the Dream Argument of step 7, as far as Descartes is concerned, shows that not even our “best” perceptions yield any certainty. For instance, one cannot be absolutely certain, merely on the basis of seeing a sheet of paper in good light from a few feet away, that there

really is a sheet of paper there; since one could have the same conscious experience in a dream. This is a skeptical position.

To understand Descartes's next step, we need to bring into focus the pattern of development that he began in step (5). This pattern is a "pro-con," or *dialectical* one. Step (5), where Descartes began his critique of the senses, was a first "con" point—one directed *against* the senses. Step (6) was a "pro" point—a point *in favor* of the senses. And step (7)—the Dream argument—is a new "con" point. This dialectical structure, in which each new point is a response to the previous one, allows Descartes to develop his critique of the senses without overlooking what can be said on their behalf.

Descartes's next point, accordingly, is a new "pro" point. But since Descartes has already reached a very skeptical position regarding the senses, this new point is a last-ditch effort to salvage something from the senses. Descartes now suggests that even if we can never be certain that we are perceiving reality rather than having a vivid dream, we can at least be sure that the images we have in our dreams are derived from reality—are "like paintings, which must have been fashioned in the likeness of things that are real" (CSM II 13, SPW 77, AT VII 19). This suggestion is essentially a hypothesis concerning the origin of dreams: their contents, though illusory, must be based upon something real that we previously perceived and thus must in some degree correspond to reality.

The details here are interesting. First, Descartes suggests that at least dream images of heads, hands, eyes, and whole bodies must be derived from those very things, somewhat as a painter's depiction of imaginary animals can only represent parts of animals jumbled up in various ways. Then, he asks himself, in effect, what if a painter comes up with something wholly fictitious and unreal? He responds that even so, at least the colors used in the painting must be real ones. He then says that likewise, even if heads, hands, eyes and so on are imaginary, the simplest and most universal elements of dream images—corporeality, extension (= three-dimensional spread-outness, or three dimensionality), shape, quantity, size, number, place, duration—must be derived from real counterparts. Descartes thus suggests that the colors used in a painting are to images of fictitious creatures in the painting as the simplest elements in dream images are to real counterparts of them. This is a very abstract analogy, since it abstracts from the fact that colors are *components* of or *ingredients* in the paintings, whereas the real things from which the simplest elements in dreams are supposed to be derived are *counterparts* of those

elements. But the point of the analogy is that both the colors and the counterparts have to be *real*: they must really exist.

From his analogy, Descartes then infers that sciences that deal with these simplest things, especially mathematics (“arithmetic and geometry”) may be more secure than those that deal with composite things. Here he is not only highlighting the apparent certainty of mathematics, but also preparing the ground for his own geometricized physics, to be discussed later in this book. In his early work, *Rules for the Direction of the Mind*, he calls the simple elements just listed “simple natures” and says that “the whole of human knowledge consists uniquely in our achieving a distinct perception of how all these simple natures contribute to the composition of other things” (CSM I 49, AT X 427).

At this point, Descartes also declares:

[W]hether I am awake or asleep, two and three added together are five, and a square has no more than four sides. It seems impossible that such transparent truths should incur any suspicion of being false. (CSM II 14, SPW 78, AT VII 21)

We shall postpone comment on this remark for a few moments, since it raises a special puzzle. So we may now enter, as the next step in our summary, the following:

8. Mustn’t the images in a dream at least be derived from something real? Further, isn’t mathematics secure even in sleep?

As already mentioned, this suggestion is a last-ditch attempt to salvage something from the senses. For it *concedes* that we can never be certain (on the basis of our senses) that we are perceiving reality, rather than dreaming, insisting only on a very minimal and indirect link between perceptual experience and reality.

Nevertheless, Descartes finds that not even this tenuous link is absolutely certain. To show why it is not, he invokes a final, extraordinary skeptical argument—perhaps the most radical skeptical consideration imaginable:

And yet firmly rooted in my mind is the long-standing opinion that there is an omnipotent God who made me the kind of creature that I am. How do I know he has not brought it about that there is no earth, no sky, no extended thing, no shape, no size, no

place, while at the same time ensuring that all these things appear to me to exist just as they do now? What is more, since I sometimes believe that others go astray in cases where they think they have the most perfect knowledge, may I not similarly go wrong every time I add two and three or count the sides of a square, or even in some simpler matter, if that is imaginable?

(CSM II 14, SPW 78, AT VII 21)

Descartes is here saying that, contrary to (8), the images in one's dreams need not even correspond to anything real. There need be no connection at all, not even the most tenuous one, between my perceptual experience and physical reality; *for perhaps there is no physical world at all!* Perhaps, instead, an omnipotent (all-powerful) God has created me such that I have experiences exactly like the ones I would have if there were such a world, that is, visual, tactile, auditory, gustatory, olfactory and kinesthetic experiences so vivid and orderly that it seems to me that I am perceiving physical things even though there really are no such things. In short, perhaps God has so created me that I hallucinate the entire physical world! How can I possibly know that this is not the case, since by hypothesis all of my perceptual experiences would be exactly the same if it were?

For the first time in the *Meditations*, then, Descartes is here calling into question the very existence of the physical world. His *Deceiver Argument*, as we may call it, goes far beyond the Dream Argument, which questioned only whether we can tell *when* we are perceiving physical things, not whether such things exist. It also goes beyond the arguments of Pyrrho, Montaigne, and all other earlier skeptics, who had suggested that the senses can deceive us about the *nature* of the physical world but never that they can deceive us about its very *existence*. It is with the Deceiver Argument, then, that Descartes implements his strategy of carrying skepticism even further than the skeptics themselves as a preparation for showing, in his subsequent *Meditations*, that even this radicalized skepticism can be refuted.

In giving his Deceiver Argument, Descartes even raises the possibility that God deceives him about simple mathematics. This is puzzling, for two reasons. First, if Descartes is going to call even simple mathematics into doubt, then his project of finding certainty seems doomed from the start; for to carry out that project, Descartes intends to use philosophical reasoning: he intends to use *logic*. But if even simple arithmetic can be doubted, then why can't logic be doubted too? Consider, for example, two

of the most obvious rules of logic, called *Modus Ponens* and *Modus Tollens*, respectively. *Modus Ponens* says that any argument of the form

If P, then Q.

P.

. . . Q.

is valid. *Modus Tollens* says that any argument of the form

If P, then Q.

not-Q.

. . . not-P.

is valid. Now these logical rules are not any simpler or easier to grasp than $2 + 3 = 5$. So if Descartes really means to doubt simple arithmetic, it seems that he must also doubt the simplest rules of logic. But then how can he legitimately *use* logic to overcome his doubt? This problem reappears in an urgent way in *Meditation III*, as we shall see in chapter 3.

The second reason why Descartes's calling mathematics into doubt in the *First Meditation* is puzzling is that throughout that *Meditation*, Descartes is examining beliefs based on the senses (as seen in point (4) of our summary). But Descartes did not believe that mathematical beliefs are based on the senses; he took them to be based on the use of reason. So why does he even mention mathematics at all in *Meditation I*? This question can also be expressed in a different terminology, namely, the *a priori/a posteriori* terminology that was made famous by Immanuel Kant (1724–1804) and is now commonplace in philosophy. An *a priori* statement is defined as one that can be known just by thinking, and such statements are said to be items of a priori knowledge. By contrast, an *a posteriori* statement (also called an “empirical” statement) is defined as one that can be known only by experience, that is, by sense perception or by introspection of one's own feelings and moods; such statements are said to be items of a posteriori, or empirical, knowledge. Using this terminology, the key point is that Descartes believed that mathematical statements are *a priori*; he did not believe that they are *a posteriori*, or empirical. For example, he would have said (as would most philosophers today) that $2 + 3 = 5$ can be known to be true just by using

one's mind—just by thinking about what the statement says. One need not consult one's experience (i.e., make any observations or perform any experiments or introspect one's own feelings or moods) to be sure that $2 + 3 = 5$ is true. This mathematical statement, like other mathematical truths, is an item of a priori knowledge, not of a posteriori, or empirical, knowledge. So our question can be put this way: Since Descartes in *Meditation I* is examining a posteriori, or empirical, knowledge, why does he even mention the a priori statements of mathematics?

A plausible solution to these puzzles has been offered by Harry Frankfurt. Drawing on remarks that Descartes made when he was questioned about the *Meditations* by a Dutch scholar named Burman, who recorded those remarks in a volume titled *Conversation with Burman*, Frankfurt suggests that throughout *Meditation I*, Descartes deliberately adopts a stance which is not really his own considered or final position, namely, that of a person who believes that *all* knowledge rests on the senses. In other words, Descartes poses as a philosophical “beginner” or “novice,” who naturally assumes that all knowledge is a posteriori and starts for the first time to reflect critically on this belief.⁶ As the argument of the *Meditation* unfolds, it becomes increasingly evident to this “novice” that his stance is untenable, because of the weaknesses of the senses that the *Meditation* itself brings to light. Thus, by the end of *Meditation I*, the novice is prepared to give up his faith in the senses and to receive the more authentic epistemology (= theory of knowledge) that Descartes will offer in his subsequent *Meditations*. The relevance of this point is that when Descartes mentions mathematics in *Meditation I*, he is still thinking of it through the eyes of the novice, who believes that mathematical knowledge, like all other knowledge, is empirical.⁷ Thus, neither the certainty regarding mathematics that Descartes expresses in his response to the Dream Argument (point 8 in our summary) nor the doubt of mathematics that he expresses in the Deceiver Argument reflect a correct understanding of mathematics as a science *not* based on sense perception but on reason. The upshot is that the doubt concerning mathematics that Descartes expresses in *Meditation I* is possible only on what will turn out to be the false assumption that mathematics is empirical.

Descartes anticipates two possible objections to his Deceiver Argument. First, it might be said that God would not allow him to be always deceived, since God is supposed to be supremely good. Descartes responds that if it

⁶ Frankfurt, *Demons, Dreamers, and Madmen*, p. 32.

⁷ Frankfurt, *Demons, Dreamers, and Madmen*, pp. 73–75.

were contrary to God's goodness to allow him to be always deceived, then it would also be contrary to God's goodness to allow him to be sometimes deceived. In order to avoid misinterpreting Descartes, it is important to understand that he is not here (or anywhere while presenting the Deceiver Argument) assuming that God exists. At this point in his *Meditations*, where he is calling his beliefs into doubt, it would obviously be illegitimate for him to make such an assumption. His point is that if God exists, then there seems to be no reason why constant deception should be inconsistent with his supreme goodness if occasional deception is not. This point raises a difficulty for anyone who believes in a supremely good God, but it does not commit Descartes himself to assuming the existence of God. Later in the *Meditations* (as we shall see), Descartes tries to prove that a supremely good God exists, and after doing so, he himself tries to solve the difficulty just raised, by arguing that human error is due to our misusing our own free-will, rather than to God's deceiving us. But at this point, he is merely showing that a simple appeal to God's goodness—unsupplemented by philosophical reflection that would explain why he allows occasional error—cannot answer the Deceiver Argument. The second possible objection that Descartes considers takes a tack almost contrary to the first: some people, Descartes says, might deny that there is a God powerful enough to always deceive. Descartes responds that if he was created by some source less powerful than God, then this only makes it more likely that he is an imperfect being who is always deceived!

Let us enter the Deceiver Argument into our summary, as a response to step (8):

9. **No. For perhaps an omnipotent God has so created me that I hallucinate the entire physical world, and even go wrong in doing simple math. If you say that a supremely good God wouldn't always deceive me, my reply is that sometimes he does allow deception: so why not always? If you deny that I was created by an omnipotent God, I answer that the less powerful was my creator, the more likely I am imperfect and so always deceived.**

The argument of *Meditation I* is now substantially complete. The position Descartes has reached is an utterly skeptical one:

I have no answer to these arguments, but am finally compelled to admit that there is not one of my former beliefs about which a

doubt may not properly be raised; and this is not a flippant or ill-considered conclusion, but is based on powerful and well thought-out reasons. (CSM II 14–15, SPW 78–79, AT VII 21)

Referring to all of the beliefs that he has so far surveyed (i.e., all beliefs that, at least from the point of view of a philosophical novice, rest on the senses), he adds: “So in future I must withhold my assent from these former beliefs just as carefully as I would from obvious falsehoods, if I want to discover any certainty” (CSM II 15, SPW 79, AT VII 22). This statement simultaneously reaffirms Descartes’s goal, reiterates his method of doubt, and indicates (by the phrase, “these former beliefs”) the full, sweeping range of beliefs that this method has now led him to withhold. Let us enter Descartes’s skeptical conclusion into our summary:

10. So far, I have found nothing that there isn’t some reason to doubt.

Before closing *Meditation I*, Descartes does one more thing. He notes that it will be difficult to stick with his decision to withhold belief from all the things that have now been found doubtful, especially since he has so long been accustomed to taking them for granted. Accordingly, in order to counterbalance his tendency to accept them, he adopts a special methodological device. He deliberately pretends that there is “a malicious demon of the utmost power and cunning [who] has employed all his energies in order to deceive me” (CSM II 15, SPW 79, AT VII 22). This evil demon does the most striking thing mentioned in the Deceiver Argument of (9): he causes Descartes to hallucinate the entire physical world, including even Descartes’s own body. We may enter this final step into our summary, as follows:

11. To be sure that I remain faithful to my resolution not to accept as true anything which isn’t absolutely certain, I shall deliberately assume that a powerful, evil demon is continually deceiving me about the existence of the entire physical world, including even my own body.

It is very important to understand that Descartes is not asserting that there actually is such a deceiver. He does not know that to be true anymore than he at this point knows that there is a (good) God. Rather, he is using the *possibility* that there might be such a deceiver as a way of enforcing his methodological doubt. Furthermore, the “deceiver hypothesis,” as

we shall call this device, will soon serve an additional, related function. In *Meditation II* it will become a kind of “litmus test” for certainty. When any given proposition, *p*, presents itself as being possibly certain (as a candidate for certainty, so to speak), Descartes will ask: “could the evil deceiver fool me about *p* (make me falsely believe that *p*)?” If the answer is *yes*, then *p* is not “certain and indubitable.” But, should there be any case where the answer is *no*, then Descartes will at last have discovered something which is absolutely certain.

3. Is the Cartesian Doubt Self-Refuting?

Although the *First Meditation* is only a few pages long, it is a seminal philosophic text that has evoked thousands of pages of commentary. Entire books have been written about particular themes in it, such as the dream argument and the deceiver argument. Obviously, we cannot possibly do justice to all of this critical commentary in a work like this one. In this section, however, we shall try to provide some assessment of the *First Meditation* by considering one possible objection to Descartes’s doubt of the senses. We shall consider this objection in relation successively to each of the three reasons that Descartes gives for doubting the senses—the deceptiveness of the senses, the Dream Argument, and the Deceiver Argument.

3.1 The Deceptiveness of the Senses

In step (5) of our summary of *Meditation I*, Descartes asserts that his senses are sometimes deceptive and concludes that he should not trust them “completely.” But how does Descartes know that his senses have sometimes deceived him? The only possible answer seems to be: by using his senses. For the only way to discover a perceptual error is by using the senses themselves. For example, how does Descartes know that the tower, which he perceived from afar to be round, is really square? Well, by subsequently getting a better view of it, from a nearer distance. So, he uses his sense of vision in order to discover and to correct his earlier error. This point has led some critics of Descartes to raise an objection. The only way a person can know that his or her senses are sometimes deceptive is by using those senses themselves. Therefore, one cannot use the premise that one’s senses are sometimes deceptive to support the conclusion that they may always be deceptive; for that very premise could never be known if the conclusion drawn from it were true. One

twentieth-century American philosopher, Maurice Mandelbaum, stated this objection as follows:

[A]n epistemologist might . . . argue that once we admit that the same tower can look round from a distance and square when seen close at hand, or that the same mountain can look one color and then another, we are no longer able to maintain that the testimony of the senses is a reliable guide to the nature of objects. His argument would be that if the senses can sometimes deceive us by virtue of giving us differing reports, it is at least theoretically possible that they always do so; or, at the least, he can challenge us to produce any clear criterion by means of which we can in every case know when our senses deceive us, and when they do not.

However, . . . one cannot prove that the senses actually *do* sometimes deceive us without assuming that they sometimes do not. I would therefore contend that this skeptical argument is self-refuting. It consists in drawing the conclusion that we can never know whether our senses are deceiving us from the fact that they actually do deceive us; however, . . . this premise—the statement that they do sometimes deceive us—could not itself be *known* to be true if the conclusion of the argument, that we can never know they are deceiving us, were itself taken as true.⁸

Although Mandelbaum does not mention Descartes by name (and perhaps would not have wished to be interpreted as specifically criticizing Descartes), this passage is an especially clear and forceful statement of an objection that is commonly made against Descartes. Drawing on the passage, we can break the objection down into three parts. First, a *self-refuting argument* is defined as an argument whose premise(s) cannot be known if its conclusion is true. Second, Descartes's reasoning in step 5 of the *First Meditation* is interpreted as follows:

- (1) If my senses sometimes deceive me, then it is possible that my senses always deceive me.
- (2) My senses sometimes deceive me.

. . . It is possible that my senses always deceive me.

⁸ Maurice Mandelbaum, *Philosophy, Science, and Sense Perception*, p. 132.

Finally, it is pointed out that this argument—although a perfectly valid case of *Modus Ponens*—is self-refuting; for if its conclusion were true, then its second premise could not be known. This is because, as we have seen, the only way to know that one's senses are sometimes deceptive is by using those senses themselves.

Before we consider whether this objection is damaging to Descartes's position, we should pause to note that skeptical arguments are, in general, vulnerable to the charge of self-refutation. Suppose that a skeptic were to advance an argument for the conclusion that nobody knows anything. Then the obvious rebuttal would be: "If nobody knows anything, then you yourself do not know anything, and so do not know the premises (or the conclusion) of your skeptical argument." This suggests that skeptics must not formulate their position dogmatically as "Nobody knows anything"; they must formulate it more cautiously. Indeed, it has sometimes been said that the only rational course for a skeptic is to remain silent—to refrain from affirming or denying anything. But this goes too far, for there are ways that skeptics can formulate their position that are not self-refuting. For example, they can say that there are reasons for suspending judgment on many matters normally assumed to be known. They can back this up by advancing arguments on both sides of a question and just leaving matters there. Or they can argue that we lack certain *kinds* of knowledge, such as perceptual knowledge, memory knowledge, or knowledge of the future. But skeptics must not try to "prove" that we don't know anything; and in giving their skeptical arguments, they must be careful not to assume that they possess the very kind of knowledge those arguments are supposed to impugn. A skeptic, then, must always walk a fine line to avoid self-refutation.

So is Descartes's reasoning in step (5) self-refuting? No, it is not; for although Descartes is frequently misinterpreted on this point, he does *not* make use of the argument (displayed on page 28 above) that is often attributed to him. He does *not* argue from the premise that his senses are sometimes deceptive to the conclusion that they may always be deceptive. Rather, he concludes that he should not trust his senses "completely." But this only means that some of his perceptions (i.e., those that occur under poor conditions of observation) should not be trusted; for as we have seen, Descartes explicitly indicates in his next step—point 6 in our summary—that he does not mean to extend this conclusion to perceptions occurring in good conditions of observation. Thus, critics who

attribute this argument to Descartes have simply not read him carefully enough. We can conclude that his reasoning in step 5 is not self-refuting.

Of course, by the time Descartes gets to the end of *Meditation I*, he does reach a thoroughly skeptical position with regard to sense perception. As one commentator has pointed out, Descartes's overall argument in *Meditation I* moves from saying that *some* perceptions are deceptive (step 5), to saying that *any* perception may be deceptive (because of the Dream Argument, step 7), to saying that *every* perception may be deceptive (because of the Deceiver Argument, step 9).⁹ Does this mean that his skepticism about the senses is self-refuting after all? To see whether it does, we must consider the Dream Argument and the Deceiver Argument.

3.2 The Dream Argument

Some Descartes scholars have offered subtle, complex reconstructions of the dream argument.¹⁰ But for our purposes, we can work with a fairly simple version of it:

1. I sometimes have vivid dreams that are qualitatively just like my “best” (waking) perceptions.
2. If I sometimes have vivid dreams that are qualitatively just like my “best” perceptions, then I cannot distinguish with certainty between my “best” perceptions and vivid dreams.
3. I cannot distinguish with certainty between my “best” perceptions and vivid dreams. [from (1) and (2)]
4. If I cannot distinguish with certainty between my “best” perceptions and vivid dreams, then even my “best” perceptions provide no certainty.
5. Even my “best” perceptions provide no certainty. [from (3) and (4)]

A paraphrase of (3) could be: “I can never be certain whether I am perceiving objects under ideal conditions of observation or having a vivid dream.” A paraphrase of (5) could be: “Even when I *am* perceiving objects under ideal conditions of observation, I cannot be certain that I *am*.”

Although this argument is valid, it is vulnerable to the charge of self-refutation. That charge could be developed as follows. How does

⁹ Bernard Williams, “Descartes’s Use of Skepticism,” p. 341.

¹⁰ See, e.g., Edwin M. Curley, *Descartes Against the Skeptics*, pp. 51–52, and Margaret Wilson, *Descartes*, pp. 17–18 and pp. 22–23.

Descartes know that (1) is true? In order to know that vivid dreams are just like his best waking perceptions, he would have to *compare* dreams with waking perceptions and note their similarity. But in order to make this comparison, he must identify his vivid dreams *as* dreams and his best waking perceptions *as* waking perceptions and then note their similarity. Now if he cannot tell his vivid dreams and his best waking perceptions apart in the first place, then he cannot identify the former as dreams and the latter as waking perceptions; so he cannot make the comparison needed to know whether they are alike or unalike. Consider an analogy. Suppose someone showed you a drawer containing one hundred one-dollar bills, informed you that fifty of the bills were counterfeit, and asked you to compare the real bills with the counterfeit ones. Further, suppose that there were absolutely no detectable difference between the genuine and the counterfeit bills (that they were indistinguishable). Then, isn't it obvious that you could not even begin the comparison? But if (3) is true, then Descartes cannot tell vivid dreams apart from waking perceptions anymore than you could tell the real bills apart from the counterfeit ones. So he cannot make the comparison needed to know that (1) is true. Thus if (3) is true, then Descartes cannot know that (1) is true. Therefore, the first step in the argument—from (1) and (2) to (3)—is self-refuting; and we need not even examine the second step, from (3) and (4) to (5).

It seems that we must concede that the Dream Argument is self-refuting, for it seems to fully satisfy the definition of a self-refuting argument. However, it can be argued on Descartes's behalf that this result is not really damaging to his position; for Descartes uses the dream argument for a special, limited purpose, namely, to show that he cannot distinguish *with certainty* between vivid dreams and his "best" perceptions and therefore that even his "best" perceptions can be *doubted*. But this use of the Dream Argument does not require that Descartes *know* that premise (1) is true, or that he be certain of its truth. It only requires that he genuinely *believe* that the premise is true. Descartes, so to speak, finds himself believing that he sometimes has vivid dreams that duplicate waking perceptions occurring under even the best conditions of observation. And this mere belief gives him a legitimate reason to doubt whether he can ever distinguish with certainty between vivid dreams and waking perceptions.

The point underlying this defense of Descartes is that a ground for doubt need not itself be something that you know or are certain of; it

need only be something that you genuinely believe. To see this, consider the following imaginary dialogue:

Q: "What time is it?"
A: (looking at his watch) "It's 9:00 P.M."
Q: "Are you absolutely certain of that?"
A: "No."
Q: "Why not?"
A: "My watch isn't always correct."
Q: "Are you absolutely certain of that?"

Notice that even if A were to answer *no* to Q's final question, A would still have his grounds for being uncertain that the time is 9:00 P.M. It would be absurd for Q to challenge A by saying: "Well then, since you are not certain that your watch is sometimes incorrect, you ought to be certain that it's 9:00 P.M." For such a challenge rests on a false assumption, namely, that a ground or reason for doubt must itself be something of which one is certain. For A to have some ground to doubt that it is 9:00 P.M., he need not be certain that his watch is sometimes incorrect. The mere fact that he genuinely believes that it is sometimes incorrect is a legitimate reason for him to be less than certain that it is now 9:00 P.M. Likewise, the mere fact that Descartes genuinely believes that he sometimes has dreams that are qualitatively just like his "best" perceptions is a legitimate reason for him to be less than certain that he can distinguish such perceptions from vivid dreams.

3.3 The Deceiver Argument

Even if what has been said so far in defense of Descartes's doubt were not wholly satisfactory, this would not matter very much; for Descartes's skepticism about the senses in *Meditation I* does not ultimately rest either on the occasional deceptiveness of the senses (step 5) or the Dream Argument (step 7). Those arguments are only preliminaries. They are merely intended to shake our faith in the senses a little, so as to prepare the way for Descartes's main skeptical argument: the Deceiver Argument. Whatever may be said about the other skeptical arguments of the *First Meditation*, the Deceiver Argument is not vulnerable to the charge of self-refutation. The basic reason for this is that the Deceiver Argument, like those other skeptical arguments, is intended to show that the *senses* provide no certainty. So if its premises could only be known to be true by

using the senses, then the argument would be self-refuting. However, the argument's premises are in fact all *a priori* statements, which can be known just by thinking. So even if its conclusion—that the senses provide no certainty—is true, this does not prevent its premises from being known. Therefore, it is not a self-refuting argument. Let us spell all this out in more detail.

To do this, we need to use the concept of an *analytic* statement. An analytic statement can be defined as a statement that is true solely in virtue of the meanings of its constituent terms. A common example is the statement, "All bachelors are unmarried." Although this statement is not worded as a definition—it does not say, "the term 'bachelor' means the same as the term 'unmarried'"—it is obviously true by definition (i.e., true solely by virtue of the meanings of the terms it contains). Sometimes analytic statements are also called "conceptual truths," since they are true solely because of the relationships between the concepts they involve. Analytic statements contrast with *synthetic* statements, which are statements whose truth or falsity does not depend solely on the meanings of terms. For example, the statement that "All bachelors are taxpayers" is synthetic.¹¹

Now an important point about analytic statements is this: if a statement is analytic, then it is *a priori*—it can be known to be true just by thinking; for to know that such a statement is true, one need only understand what it says. One need not consult experience (make any observations or perform any experiments or introspect one's feelings). Of course, in order to learn the meanings of words, beings constituted, or "wired," like humans do need to have various sorts of experiences. At first, this fact might seem to conflict with saying that analytic statements—or indeed any statements—are *a priori*. But to see better why a statement like "All bachelors are unmarried" is *a priori*, compare it again with "All bachelors are taxpayers." Even after one knows what the latter statement means, one can still be totally in the dark as to whether it is true or false; for its truth or falsity depends on facts about law and society that can only be known by experience. On the other hand, once one has learned what "All bachelors are unmarried" means—which admittedly requires various sorts of experience—no further experience is needed to

¹¹ Some contemporary philosophers, notably the American thinker W. V. Quine (1908–2000), reject the distinction between analytic and synthetic statements. This discussion assumes that the distinction is a tenable one.

know that this statement is true. Thus the statement can be known to be true independently of any experience, except for the experience(s) needed to learn the meanings of its constituent terms. And this is all that is meant by saying that it can be known just by thinking, or is an *a priori* statement.

Having defined the concept of an analytic statement, and noted the principle that analytic statements are *a priori*, we can begin to see why the premises of the Deceiver Argument are *a priori*. It is because the argument turns mainly on a purely analytic proposition, knowledge of which is accordingly *a priori*. This proposition is the causal conception of perception (CCP):

(CCP) For any person S and material object M , S perceives M at time t only if M is a cause of S 's perceptual experience at t .

According to this conception, which underlies Descartes's discussion of perception throughout the *Meditations* and which is accepted by many contemporary philosophers, as well, perception by definition involves a causal element: the object perceived must be one of the causes of the perceiver's experience. For example, if S now sees a pen, then the pen must be one of the causes of S 's visual experience. The pen is, of course, not the only cause: the causes of S 's visual experience also include light striking S 's retinae, processes in S 's nerves and eyes, and so forth. But the pen must be a cause, or one of the causes, of S 's present visual experience. This is why CCP says that M is *a cause*, rather than *the cause*, of S 's present perceptual experience.

Before showing exactly how the Deceiver Argument rests on CCP, we should support our claim that CCP is an analytic truth. One way to do so is to suppose that S were to say

(A) I see a pen, but it is not the case that a pen is one of the causes of my present visual experience.

Surely, A is a contradiction: if the pen is not even one among the causes of my present visual experience, then it is absurd to say that I see it. The first four words of A assert that I see the pen; and the remaining words deny, or "take back," this assertion. But the negation of a contradiction is always an analytic statement. For example, the negation of "Some bachelors are married" (which is a contradiction since it means that at least

one individual is both married and not married) is “All bachelors are unmarried,” which, as we saw, is analytic. Now the negation of A is

- (B) If I see a pen, then a pen is a cause of my present visual experience.

One way to see why the negation of A is B is to note that the logical form of A is *p and not-q* (e.g., “John is wearing his right shoe and not wearing his left shoe”). The denial of this form is *not (p and not-q)* (e.g., “it is not the case that John is wearing his right shoe and not wearing his left shoe”). But the latter is equivalent to *if p, then q* (e.g., “If John is wearing his right shoe, then he is wearing his left shoe”), which is the logical form of B.] Thus, B is analytic. Now B means exactly the same thing as

- (C) I see a pen only if a pen is a cause of my present visual experience.

This is because the logical form of B, *if p then q*, is equivalent to the logical form of C, *p only if q*. (To see this, note that “If my car runs, then there is gas in the tank” says exactly the same thing as “My car runs only if there is gas in the tank.”) So, since B is analytic and says the same thing as C, C is analytic. But if C is analytic, then it seems evident that analogues of C involving senses other than vision and persons other than “I” (e.g., “Mary smells a rose only if a rose is one of the causes of her present olfactory experience,” “John tastes a potato chip only if a potato chip is one of the causes of his present gustatory experience,” “Henry hears a bell only if a bell is one of the causes of his present auditory experience,” etc.) are also analytic. But then CCP, which is nothing but the general statement that summarizes all such particular ones, is analytic.¹²

A different demonstration that CCP is analytic has been offered by the philosopher Paul Grice (1913–1988). Grice’s demonstration goes as follows. Suppose that S is having a visual experience exactly like seeing a clock on the shelf—or, as Grice puts it, that it looks to S as if there is a clock on the shelf. Furthermore, suppose that there really is a clock on the shelf within S’s field of view, before his eyes. Is this sufficient for it

¹² A fuller presentation of this argument can be found in my *Perceptual Knowledge*, pp. 81–82.

to be true that *S* sees the clock? No, it is not; for suppose, as is logically possible, that *S*'s visual experience is being produced by an expert directly stimulating *S*'s cortex or by some kind of post-hypnotic suggestion, so that even if the clock's position on the shelf were altered or the clock were entirely removed, *S*'s visual experience would remain unchanged: it would continue to look to *S* as if there is a clock on the shelf. In that case *S* does not *see* the clock, even though it is there before his very eyes. And the reason he does not see it is that it plays no part in causing his visual experience.¹³ Grice's example shows very clearly that in order for *S* to see the clock, it is not enough that (1) *S* has a visual experience in which it looks to him as if there is a clock on the shelf and (2) there actually is a clock there in front of *S*'s eyes. Rather, *S* does not see the clock unless (3) the clock is also a cause of *S*'s present visual experience. The concept or definition of seeing thus contains an inexpungible causal element. It seems safe to generalize from what is here true of vision to perception in general, and so to conclude that CCP is a conceptual, analytic truth about the nature of perception. (More precisely, CCP expresses a logically necessary condition, and so a partial definition, of perception.)

How does CCP relate to the deceiver argument? Well, the basic point made by the argument is that any perceptual experience that *M* causes *S* to have might be exactly duplicated by God, or (to switch to the possibility envisioned in Descartes's deceiver hypothesis) by some powerful, evil demon. Therefore, *S* can never be certain that *M* is causing the experience, and hence, given CCP, can never be certain that she is *perceiving M*. Let us make this argument even more explicit. From CCP, it follows that:

- (1) I can sometimes be certain that I perceive a material object *M* only if I can sometimes be certain that *M* is causing my perceptual experience.

Now we may also assert that

- (2) I can sometimes be certain that *M* is causing my perceptual experience only if it is not the case that any (every) perceptual experience caused by *M* could be caused in some other manner.

¹³ H. P. Grice, "The Causal Theory of Perception," p. 461.

But the key idea of the deceiver argument is that

- (3) Any (every) perceptual experience caused by *M* could be caused in some other manner (e.g., by an evil demon).

It follows from (1)-(3), however, that

- (4) I can never be certain that I perceive *M*.

The conclusion, (4), follows from premises (1)-(3), because the argument has the valid form

$$\begin{array}{c} p \text{ only if } q. \\ q \text{ only if not-}r. \\ r. \\ \hline \therefore \text{not-}p. \end{array}$$

Probably, you can see by reflecting for a moment that this form is valid. Another way to see this is to note again that *p only if q* says the same thing as *if p, then q*. So, the above form is equivalent to this one:

$$\begin{array}{c} \text{If } p \text{ then } q. \\ \text{If } q \text{ then not-}r. \\ r. \\ \hline \therefore \text{not-}p. \end{array}$$

But this last form is valid: from *if q then not-r* and *r*, one can deduce *not-q* by *Modus Tollens* (and Double Negation); and then from *not-q* and *if p then q* one can deduce the conclusion, *not-p*, by *Modus Tollens*.

Not only is the Deceiver Argument valid, but we can now see more fully that it is not vulnerable to the charge of self-refutation. For premise 1, following as it does from CCP, is itself analytic and so a priori. (This is because if P is analytic and Q follows logically from P, then Q is analytic too: analyticity is hereditary with respect to entailment. Strictly speaking, (1) does not follow from CCP alone. Rather, it follows by *Modus Ponens* from CCP together with *If CCP, then (1)*. But this last statement is itself a long analytic statement, as can be seen by substituting for “CCP” and for “(1)” the clauses that these labels abbreviate.) Premise 2, we may assert,

is also analytic and so a priori: it depends for its truth solely on the meanings of its constituent terms, notably the term “certain.” Furthermore, premise 3 expresses merely a logical possibility; and so our knowledge of it does not depend on the senses either. Finally, it would be difficult to maintain that the sorts of experiences required to learn the meanings of the terms in (1), (2), and (3) must be genuine perceptions of reality. It seems that even if, as in the evil-demon scenario, all sense experiences were hallucinatory, there would be no reason in principle why one could not learn the meanings of these terms.¹⁴ Consequently, the truth of the argument’s conclusion would not prevent the premises from being both understood and known. Therefore, the argument is not self-refuting.

The deceiver argument still haunts the pages of contemporary books and articles on epistemology, though nowadays it is usually put in a more “scientific,” modern-sounding way. Any perceptual experience caused by a material object stimulating one’s sense-receptors (eyes, ears, nose, etc.), it is argued, might instead be caused by a very advanced neurophysiologist (or team of neurophysiologists) directly stimulating one’s brain with painless electrodes. Perhaps, it is then suggested, all of our perceptual experience is caused in some such way, so that we never really perceive material objects at all, but only hallucinate them. How can we possibly know that this is not so, since our perceptual experience would be exactly the same if it were so? It is not hard to recognize this line of reasoning as being Descartes’s Deceiver Argument in modern dress.

As we shall see when we come to the *Sixth Meditation*, Descartes himself tried to refute the deceiver argument. He thereby hoped to answer skepticism once and for all by refuting the most radical argument in its favor—an argument he had himself invented. To measure his success, however, we must first turn to *Meditation II*, where Descartes takes the first step toward answering the deep, unsettling skepticism generated in *Meditation I*.

¹⁴ Some contemporary philosophers would object to this claim, on the ground that it assumes the possibility of a “private language” and that this assumption was proved wrong by Ludwig Wittgenstein (1889–1951) in his *Philosophical Investigations*. There is no consensus among philosophers, however, whether Wittgenstein’s argument against the possibility of a private language is sound; nor is there even agreement concerning the exact nature of that argument.

Meditation II

The Cogito and the Self

1. Descartes's "I am thinking, therefore I exist"

In the opening paragraph of *Meditation II*, Descartes reminds himself of the deeply troubling doubts he raised in *Meditation I*. To convey how unsettling these doubts are, he uses an image. He compares himself to a man caught in a deep whirlpool, who is so disoriented that he can neither stand on the bottom nor swim to the surface. Nevertheless, he says that he will persist in his quest for certainty, by continuing to withhold assent from everything that is doubtful, in the hope of finding just one thing that is certain. Shifting to a different image, he compares himself to Archimedes, the Greek mathematician who boasted that if he had a long-enough lever mounted on one absolutely fixed and immovable point, he would be able to move the entire earth. Likewise, Descartes hopes to find just one absolutely certain piece of knowledge, so as to build an entire system of knowledge upon it. The shift from the image of the man tumbled around in the whirlpool to the image of Archimedes symbolizes very well the path that Descartes will traverse in the *Meditations*, from total doubt to certainty about many things. We may begin a summary of *Meditation II*, then, as follows:

1. I shall continue to withhold assent from whatever is even slightly doubtful, just as though it were false, until I find just one thing that is certain.

Accordingly, Descartes reminds himself of what is in doubt. Perhaps everything that he seems to see or remember is unreal. Perhaps "body,

shape, extension, movement and place" (i.e., the entire material world) are only illusions. Perhaps the sole certainty is that there is no certainty to be had about the world. Nor can Descartes be certain that a God, or any other being, has put these very doubts or thoughts into his mind. For perhaps he produces them himself. But then, how about this self? Can its existence be doubted too? To be sure, Descartes can and does doubt the existence of his *body*. But, he now asks, does this mean that he can doubt that *he exists*? This question has brought Descartes to the very brink of discovery. So let us record it in our summary:

2. I have convinced myself that perhaps the entire physical world, including my own body, does not exist. Does it follow that perhaps I do not exist?

Now comes Descartes's answer: "No: if I convinced myself of something (or thought anything at all [added by Descartes in the French translation]), then I certainly existed" (CSM II 17, SPW 80, AT VII 25). Here then, at last, is the first certainty that he was seeking: it is his own existence, as revealed in his very attempt to doubt it. Let us record this moment of discovery in our summary:

3. No: if I convinced myself of something, or even thought anything at all, then, certainly, I existed.

Descartes immediately applies to this first discovery his "litmus test" for certainty, namely the evil demon hypothesis of *Meditation I*—"But there is a deceiver of supreme power and cunning, who is constantly deceiving me." Of course, the words "there is" here mean "suppose there is." As before, Descartes is not asserting that there is a deceiver; rather, he is using the possibility that there might be one to test the certainty of "I exist." And the result is that this statement unquestionably meets the test; for how can the demon deceive Descartes unless Descartes exists? So Descartes concludes that "I exist" is absolutely certain. To paraphrase the test and its result:

4. But what if there is an evil demon constantly deceiving me? Then I certainly exist, if he is deceiving me; and he can never cause me not to exist so long as I think that I exist. So "I am, I exist" must be true whenever I assert it or think it.

Steps 3 and 4 contain Descartes's famous proof of his own existence, called the *cogito*. This Latin name comes from the formulation *Cogito, ergo sum*, which means "I am thinking, therefore I exist." This argument, which was also expressed by Descartes in French as *Je pense, donc je suis*, is usually rendered in English as "I think, therefore I am (or exist)." But "I am thinking, therefore I exist" is arguably a better translation, for a reason offered by John Cottingham:

[T]he correct English translation of *cogito/je pense*, when these words occur in Descartes' discussion of the certainty of his existence, should employ the so-called continuous present—'I am thinking'—rather than the simple present, 'I think.' For what makes me certain of my existence is not some static or timeless fact about me—that I am one who thinks; rather, it is the fact that I am at this moment engaged in thinking. And so long as I continue to be so engaged, my existence is guaranteed.¹

Interestingly enough, Descartes never uses the sentence "I am thinking, therefore I exist" in the *Meditations*. Furthermore, not everything that Descartes says in steps 3 and 4 turns on the idea that he exists just *because* he is thinking. For example, in step 4 we find the idea that to be deceived he would have to exist. This is based on the general principle that to be the *object* of some act (such as an act of deception) one must exist. This principle has nothing special to do with the connection between thinking and existing: another application of it might be that in order to be punched in the nose, one must exist. Still, most of what Descartes says about the certainty of his own existence, both in the *Meditations* and in his other works, does turn on the idea that his thinking proves his existence. Thus, in step 3, we can find the arguments "I convinced myself of something, therefore I existed" and "I thought of something, therefore I existed"; while in step 4 we can find the arguments "I now think that I exist, therefore I cannot now be caused not to exist" and "I think at time *t* that I exist, therefore 'I exist' is true at time *t*." Surely these arguments are variations on Descartes's fundamental idea: that his thinking proves his existence. As for the famous sentence itself, one place where Descartes does use it is in his *Discourse on the Method*, where he says:

¹ John Cottingham, *Descartes*, p. 36.

I noticed that while I was thus trying to think everything false, it was necessary that I, who was thinking this, was something. And observing that this truth '*I am thinking, therefore I exist*' was so firm and sure that all the most extravagant suppositions of the sceptics were incapable of shaking it, I decided that I could accept it without scruple as the first principle of the philosophy I was seeking. (CSM I 127, SPW 36, AT VI 32; see also CSM I 195–6, SPW 162–3, AT VIIIA 7–8)

Since most of what Descartes says to show that his existence is certain does turn on the connection between his thinking and his existence, we shall restrict our attention from now on to the classic formulation of the *cogito* as "I am thinking, therefore I exist."

2. The Certainty of One's Own Thoughts

Let's begin with a question that was raised by Pierre Gassendi, the author of the fifth set of *Objections* to the *Meditations*. (Descartes published the *Meditations* together with seven sets of objections written by other philosophers and theologians of his day, and with Descartes's own replies to these. This entire debate, called the *Objections and Replies*, is a fascinating one, which we shall draw upon frequently in what follows.) Gassendi said: "But I do not see that you needed all this apparatus, when on other grounds you were certain . . . that you existed. You could have made the same inference from any one of your other actions, since it is known by the natural light that whatever acts exists" (CSM II 180, M 68, SPW 126, AT VII 259).

Descartes's reply was this:

When you say that I "could have made the same inference from any one of my other actions" you are far from the truth, since I am not wholly certain of any of my actions, with the sole exception of thought. . . . I may not, for example, make the inference "I am walking, therefore I exist," except in so far as the awareness of walking is a thought. The inference is certain only if applied to this awareness, and not to the movement of the body which sometimes—in the case of dreams—is not occurring at all, despite the fact that I seem to myself to be walking. Hence from

the fact that I think I am walking I can very well infer the existence of a mind which has this thought, but not the existence of a body that walks. And the same applies in other cases. (CSM II 244, M 68, SPW 126–7, AT VII 352)

Descartes is here making at least two points: (a) “I am walking,” and other reports of my physical actions, are not certain; (b) “I am thinking,” and more specific reports of my own thoughts (such as “I think that I am walking”), are certain. This is why “I am thinking,” but not “I am walking,” can be used as a premise from which to prove my existence.

Point (a) is easy enough to understand, in light of the doubt about everything physical raised in *Meditation I*. Point (b), on the other hand, takes us beyond the skeptical doubts of *Meditation I* and introduces us to one of Descartes’s most important and influential positive doctrines. This is that each of us has absolutely certain, indubitable knowledge of his/her own present thoughts. Descartes’s doctrine is not merely that “I am thinking” is certain for each of us; it is much broader than that. The doctrine is that *all* beliefs, assertions or judgments *about one’s own present thoughts* enjoy a special certainty that makes them immune to the skeptical doubts of *Meditation I*. We may formulate it like this. Let p be any statement that you believe. Then even if p itself is uncertain, “I am thinking about p ” is certain, as is “I believe that p .” For example, even if “there is a horse in the field” is uncertain, “I am thinking about there being a horse in the field” and “I believe that there is a horse in the field” are certain. Thus for every statement that you think about or believe, whether it is certain or not, there is a corresponding one (or more) which is certain. Also, let p be any statement of the type: “I perceive X” (where “perceive” means “see,” “touch,” “hear,” “taste,” “smell”). Then even if p is uncertain, “I seem to perceive X” is certain. This will yield as many certain statements as there are things you seem to perceive. For example, “I seem to see a horse,” “I seem to touch a hand,” “I seem to hear a car,” “I seem to smell a rose,” “I seem to taste a pear,” and so on are all statements that can be certain. Still other types of statements that can be certain, according to Descartes’s doctrine, are:

I doubt that p . (e.g., “I doubt that bodies exist.”)

I judge that p . (e.g., “I judge that this wax is hard.”)

I want X. (e.g., “I want to know more.”)

I don’t want X. (e.g., “I don’t want to be deceived.”)

I imagine O. (e.g., “I imagine a unicorn.”)

What all these and similar statements have in common is that they describe only one's own present state of mind; they do not make any claim about anything existing independently of one's own thinking. (Notice that this is what distinguishes "I *seem* to see a horse" from "I *see* a horse": "I seem to see a horse" is to "I see a horse" as "I think I am walking" is to "I am walking.") Descartes calls the subject matter of such statements *cogitationes* (Latin for "thoughts," sg. *cogitatio*); his view is that *cogitationes* constitute an easily overlooked but crucially important area of certainty. For this reason, he is prepared to substitute any *cogitatio*-statement (i.e., any statement about one of his own present thoughts, like the statements just listed) for the premise "I am thinking" in the *cogito*. For example, toward the end of *Meditation II* he uses the argument, "I judge that a piece of wax exists, therefore I exist" (CSM II 22, SPW 86, AT VII 33); and in *The Search for Truth* he uses "I doubt, therefore I exist." (CSM II 410, AT X 515). Descartes's view that all *cogitatio*-statements are certain explains why he formulates the *cogito*'s premise in such a large variety of different ways, both in his *Meditations* and in his other works.

How plausible is Descartes's view? To answer this question, let us evaluate Descartes's view by using his own test of certainty. To do this, we can ask the following sequence of questions:

Could an evil deceiver make it the case that

1. I think that there is a physical world, but there is no physical world?
2. I think that I have a body, but I do not have a body?
3. I think that I am walking, but I am not walking?
4. I think that I am thinking, but I am not thinking?
5. I think that I see a horse, but I do not see a horse?
6. I think that I seem to see a horse, but I do not seem to see a horse?

It should be obvious, in light of previous discussion, that the answer to questions (1), (2), (3), and (5) is *yes*. Notice, however, that the answer to question (4) has to be *no*. The deceiver cannot possibly make me think that I am thinking when I am not thinking, for to think that one is thinking *is* to be thinking! Thus, "I am thinking" passes Descartes's "litmus test" for certainty. Indeed, even without using the device of the deceiver, we can see that Descartes has made a valid point: One cannot possibly think *falsely* that one is thinking, since to think that one is

thinking is already to be thinking. Thus, “I am thinking” is absolutely indubitable. The point is exceptionally well put by Fred Feldman:

Let’s consider the . . . premise, “I think” . . . Why should we view this as being . . . certain for Descartes? The answer, I believe, turns on a remarkable feature of thought. If you think you are walking, you might be mistaken. Maybe you are only dreaming that you are walking. However, if you think you are thinking, you must be right. You cannot think that you are thinking, if you are not thinking. So if an evil demon causes you to think that you are thinking, he won’t cause you to make a mistake. You will be right. You will be thinking. If, in a dream, you think that you are thinking, you will be right. There is no possibility of error here. You cannot make a mistake, if you think you are thinking.²

Notice that the point here is not that the evil deceiver could not make you think that you are not thinking; no doubt he could do that—he could make you gullible enough to think that you are not thinking when in fact you are thinking. What he could not do is make it *false* that you are thinking when you think that you *are* thinking.

What about question (6)? Here matters are more complicated. To see this, suppose that the evil deceiver confuses me about the difference between horses and zebras: he makes me think that horses, rather than zebras, have stripes. Now, suppose that at a certain time, I have an experience as of seeing a striped, equine animal. (The phrase “experience *as of* seeing” is here meant to indicate that the experience need not be an actual seeing: it might be an actual seeing, but it might also be a dream or a hallucination, it does not matter which.) Is this a case where I wrongly think that I seem to see a horse, so that the answer to (6) is yes? To answer this question, we need to note that (6) is ambiguous, because “seems” has more than one meaning. It can signify *belief*, as when the doctor says, “you seem to have an infection.” Here the doctor is using “seems” to express her belief that the patient has an infection. The word is also used to express belief in such locutions as “It seems to me that . . .,” “It would seem that . . .,” and “It seems that . . .” But “seem” can also be used in a very different way; namely, to signify the quality of an experience; as in “The moon seems flat and yellow tonight.” Here the speaker is not

² Fred Feldman, *A Cartesian Introduction to Philosophy*.

expressing a belief that the moon really is flat and yellow. Rather, the speaker is only describing how the moon looks. This ambiguity of “seem” carries over to the statement “I seem to see a horse.” This statement could mean (a) “I believe that I see a horse”; or it could mean (b) “I have an experience as of seeing a horse.” So question (6) could be a way of asking,

- (6a) Could it be true that I think that I believe that I see a horse, though I don’t believe that I see a horse?

Or it could be a way of asking,

- (6b) Could it be true that I think that I have an experience as of seeing a horse, though I don’t have an experience as of seeing a horse?

Now it seems that the answer to (6a) is *no*; for even if the animal that I think I believe I see is striped, it is still true that I believe it to be a horse. Just because I falsely believe that horses are striped, it does not follow that I am not in a state of believing that I see a horse. On the other hand, the answer to (6b) seems to be *yes*; for if I am having an experience as of seeing a striped animal, then I am wrong in thinking that it is an experience as of seeing a horse. It is really an experience as of seeing a zebra. Perhaps, however, we can revise (6b) so that the answer to it will again be *no*, as follows:

- (6c) Could it be true that I think that I have an experience as of seeing what I *take* to be a horse, though I do not have an experience as of seeing what I take to be a horse?

At this point, someone might object that if the deceiver were powerful enough, he could cause me to go wrong in a different way. He might make it true that I think I believe I see a horse, even though I really *disbelieve* that I see a horse; or that I think I have an experience as of seeing what I take to be a horse, even though I really have an experience as of seeing what I take *not* to be a horse. Now it is not clear that such states of mind as these are possible. But even if they are, we can defend Descartes against this objection by using a point made in the previous chapter. There we argued, you will recall, that it was legitimate for Descartes to dismiss the possibility that he might be insane. But a person who thought he believed he saw a horse while really disbelieving this, or who thought

he had an experience as of seeing what he took to be horse while having an experience as of seeing what he took not to be a horse, would be so radically confused about his own conscious beliefs as not even to be sane. So, Descartes can legitimately assume that he is not in such a condition. His basic question is not, “How can an insane person use philosophical reasoning to discover certainty?” Rather, he is asking how a rational person can use philosophical reasoning to arrive at certainty. It is already built into this question that the person is not radically confused about her own conscious beliefs. Descartes’s test for certainty—“could the deceiver fool me about *p*?”—must be understood to include the stipulation that the person signified by “me” is sane.

It appears, then, that Descartes’s view about the certainty of *cogitatio*-statements is plausible. The statement “I am thinking” can be absolutely certain. And although this is a delicate matter about which more could be said, it appears that many other (though perhaps not all) *cogitatio*-statements can be certain too, provided they are carefully formulated.

3. A Problem for the *Cogito*

So far we have confined our discussion of the cogito to its premise, “I am thinking.” The next matter to be considered is the inference from this premise to the conclusion, “I exist.” At first this inference looks obvious and unproblematic. But it will prove to be as perplexing upon reflection as it was compelling at first glance.

The basic difficulty can be put very simply. What entitles Descartes to use the first-person pronoun “I” in the premise of his proof? Descartes’s use of this pronoun seems to assume or presuppose the very thing he is supposedly proving; namely, that he exists. Thus, the *cogito* seems to be what is called a *question-begging argument*, that is, according to the standard definition of “question-begging,” an argument that takes for granted or assumes the very conclusion that it is supposed to prove.

Bertrand Russell (1872–1970), one of the twentieth century’s greatest philosophers, makes a suggestion that might seem to solve the difficulty. He suggests that the “I” in “I am thinking” ought to be regarded only as a grammatical convenience, rather than as referring to a self or person who exists as something distinct from the thinking. After all, it would not have been grammatical for Descartes to express the *cogito*’s premise merely as “thinking” or as “am thinking”: the “I” is needed to obtain a

grammatical sentence. But the “I” need not therefore refer to something that exists over and above the thinking any more than the pronoun “it” in the sentence “it is raining” refers to something distinct from the rain. Indeed, that “it” could be eliminated without changing the sentence’s meaning, by saying “Rain is falling.” Likewise, Russell suggests, the premise of the *cogito*, strictly speaking, asserts only “There is a thought” or “Thinking is occurring now.”³ If this premise is substituted for “I am thinking,” then the argument no longer begs the question.

Unfortunately, Russell’s suggestion does not solve the difficulty; for it is certainly not valid to argue, “there is a thought, therefore I exist.” For, supposing I had ceased to exist but *someone else* were thinking, the premise would be true and the conclusion false. His suggestion, if taken as an attempt to defend the *cogito*, goes from the frying pan into the fire.⁴ Nor would it help to add the premise “If there is a thought, then I exist,” for the same supposition shows that this premise is false.

The problem that the *cogito* seems to run into, then, is this. If we stick with Descartes’s famous formulation, “I am thinking, therefore I exist,” then the argument appears to be question-begging, whereas if we substitute the Russellian formulation, “There is a thought, therefore I exist,” then the argument is invalid. Thus, the *cogito* appears to be either question-begging or invalid.

Some philosophers, motivated by this problem or variations on it, have tried to interpret the *cogito* as something other than an argument for one’s existence. For example, in 1962 the Finnish philosopher Jaakko Hintikka, in a famous article in *The Philosophical Review* entitled “*Cogito, ergo sum: Inference or Performance?*” argued for what he called the *performatory* rather than *inferential* interpretation of the *cogito*. Hintikka’s basic idea was this. Suppose that you try to doubt your own existence. You will immediately discover that this is impossible, because in the very attempt to think that perhaps you do not exist, your existence is manifest. Furthermore, your certainty that you exist is not then based on any *argument*; rather, it is based solely on the act or performance of trying to doubt your existence—an act that causes you to see that you exist, almost as playing music causes sound. The relation between the thinking in the *cogito* and the realization that I certainly exist is not one between premise

³ Bertrand Russell, *A History of Western Philosophy*, p. 567. Russell’s exact formulation is “There are thoughts.”

⁴ I do not mean to imply that Russell himself offers his suggestion as a defense of the *cogito*, but only to fault the idea that it might successfully serve as such a defense.

and conclusion, but rather between process and product, so that the notion of question-beggingness cannot even be applied.

This novel interpretation of the *cogito*, however, has been generally rejected by Descartes scholars, for two different reasons. First, it makes the certainty that I exist depend narrowly on only one specific thought; namely the attempt to doubt my own existence. But, as we have seen, Descartes believed that *any* of his thoughts, regardless of its content, established his existence. Second, when one asks exactly *why* trying to doubt my existence causes me to be certain that I exist, the only clear answer seems to be that I accept the *argument*: “If I try to doubt my existence, then I exist; I am trying to doubt my existence; therefore I exist.” But then the performativey *cogito* reduces to an inference or argument after all. Accordingly, it will be best for us to pursue the topic of the *cogito* by considering why Descartes thought that “I am thinking, therefore I exist” is a successful *argument* for, or *proof* of, one’s own existence. Such an approach has the further advantage that it will bring out several underlying principles of Descartes’s thought.

4. The Substance Theory

A key underlying principle that Descartes appeals to when explaining the *cogito* to his critics is what I shall call “the substance theory.” This section will temporarily digress from the *cogito* in order to expound the substance theory and the main argument for it, so that we can then see, in the following section, how this theory can be used to support the *cogito*.

The substance theory is an answer to the philosophical question: *What is a thing?* In order to understand this rather strange-sounding question, we must first understand a distinction that philosophers make, between a *thing* and a *property*. To appreciate the need for this distinction, reflect on the following, fallacious argument.

- (1) Socrates is Socrates.
- (2) Socrates is snub-nosed.
- (3) Socrates is fat.
- (4) Socrates is wise.

∴ One thing (Socrates) is many things.

The conclusion is absurd: a single thing cannot possibly be identical with many different things. What then is wrong with the argument? The

answer is that the “is” in (2), (3), and (4) is being understood in the same sense as the “is” in (1). But this is an error. For the “is” in (1) is what logicians call an “is of identity”; while the “is” in (2)–(4) is an “is of predication.” The “is” of identity is used to assert that a designated item is one and the same entity as some item, as for example in “The Morning Star (Venus seen in the morning) is the Evening Star” (Venus seen in the evening). The “is” of predication is used to attribute a characteristic or property to a thing, as in “Venus is round.” To see this distinction more clearly, notice that we could substitute the sign “=” for the “is” in (1); while it would be wrong to do so in (2)–(4).

Corresponding to the logical distinction between these two “is’s,” is a fundamental distinction within reality: that between a thing and a characteristic or property. “Socrates” designates a thing—an animate, living one. But “snub-nosed,” “fat,” and “wise” do not designate things; they designate properties. What the fallacious argument about Socrates shows is that in order to avoid the “one-many” paradox generated in its conclusion, we must include properties as well as things in our ontology. (The word “ontology” means “an account of what there is.” It is derived from *logos*, Greek for “account,” and *onto*, Greek for “being.”)

Having made the ontological distinction between a thing and a property, we can return to the question: what is a thing? There are two traditional, opposed answers to this question. The *bundle theory* holds that a thing is merely a collection of coexisting properties. For example an apple, according to the bundle theory, is nothing but roundness, redness, tartness, squashiness, and so on coexisting at a certain place and time. The *substance theory* holds that a thing is composed of various properties *plus* an underlying substance to which these properties belong. The apple, on this view, is composed not just of the properties just mentioned, but also of an underlying substance in which all these properties are said to “inhere.” The bundle theory is favored by Empiricist philosophers, such as Berkeley and Hume in the eighteenth century and Bertrand Russell in the twentieth century. The substance theory was upheld by Aristotle and most medieval thinkers, and in the Modern period by the Rationalists Descartes, Spinoza, and Leibniz. Since the substance theory is the richer one—postulating “more” in a thing than does the bundle theory—let us inquire into the rationale for it. Why not adopt the bundle theory’s simpler view that a thing is just a collection of coexisting properties?

The most salient answer is provided by an argument that we shall call the *argument from change*. This argument, which dates back at least to Aristotle,

can be discerned in Descartes's famous illustration in *Meditation II* of the piece of wax. Suppose, Descartes says, that I heat a piece of wax that has been freshly brought from a beehive. As I heat the wax, its properties change: its hardness is replaced by a soft, gooey texture, its lumpish shape by an elongated one, its brown color by a translucent tint, its fragrant aroma by a smoky smell. Even its capacity to make a tapping noise when struck with a finger is lost. Yet one and the same piece of wax still exists despite all these changes. Why is this so? Why is it not the case, instead, that the wax ceases to exist, and that another, new object begins to exist? The answer proposed by the argument from change is that although the properties of the wax have changed, the underlying substance has not: one and the same *substance* still exists, and continued to exist throughout the process of change.

The argument from change is an important one, so let us try to formulate it carefully. In order to do this, we must first make a distinction between a *determinate* and a *determinable* property. A determinate property is one that is absolutely specific, and a determinable property one that is not. Consider, for example, the property of being elliptical ("ellipticalness"). This is a determinable, rather than a determinate, property, because there are many different elliptical shapes: highly elongated ones, moderately elongated ones, and so on. A determinate property must be absolutely specific. So, only an ellipse satisfying a particular equation would be a determinate property. Color is another example of a determinable property, since there are many different shades of color. Even particular colors like red, blue, and yellow are determinables rather than determinates, since there are many shades of each of these colors. Only absolutely specific shades of a color are determinate properties. (It does not matter that we may not have a name for each such shade.)

Having distinguished between determinate and determinable properties, we can state the argument from change in proper form:

- (1) We can distinguish between (a) all of a thing's determinate properties changing without the thing's ceasing to exist, and (b) a thing's ceasing to exist.
- (2) We can distinguish between (a) and (b) only if a thing is composed, in addition to its properties, of a permanent, underlying substance.

∴ A thing is composed, in addition to its properties, of a permanent, underlying substance.

To grasp this argument, think again of Descartes's example of the melting wax. The example is supposed to illustrate a case of (a), in contrast to (b). Premise (2) can be seen as a reply to the challenge, What justifies us in taking it to be a case of (a) rather than (b)—that is, in thinking that the wax, all of whose determinate properties have changed, *hasn't* therefore ceased to exist? The conclusion, which follows validly from the premises, is a statement of the substance theory itself. Thus, the argument purports to demonstrate that the substance theory is correct.

You may be wondering why the term "determinate" is needed in premise 1. The answer has to do with the nature of change. We say, of course, that when a thing alters, its properties or characteristics have changed. But what does this mean? Not that the properties *themselves* have changed, because a property itself can never change: red, for example, is just red, and to say that it had changed would really mean that it had been *replaced* by another property, say by the property blue. Thus, to say that a thing's properties have changed is to say that it has lost certain properties and acquired others. Suppose, then, that the word "determinate" were left out of premise 1. Then the premise would say that a thing could lose all of its properties, *including all its determinable properties*, without ceasing to exist. But this would be false; for a thing's determinable properties include shape and size—not this or that specific shape and size, but just having some shape and size or other, or having what Descartes called "extension"—and a (physical) thing cannot lose all shape and size whatsoever without thereby also ceasing to exist. Thus, the term "determinate" must be included in premise 1 for that premise to have a chance of being true.

You may now ask: why then isn't the word "determinate" also included in premise 2? The answer is that to preserve the argument's validity, the conclusion would then also have to include that word, and so would have to read:

A thing is composed, in addition to its determinate properties, of a permanent, underlying substance.

But this is false; for it means that a thing is composed *only* of its determinate properties and its underlying substance. But a determinate property cannot possibly be present unless a corresponding determinable property is also present; for example, squareness cannot be present unless shape is present, and redness cannot be present unless color

is present. Thus, in order for the argument's conclusion to stand a chance of being true, the word "determinate" must *not* be included in premise 2.

Before seeing how the substance theory can be used to support Descartes's *cogito*, we must be sure to understand two further points. First, in his discussion of the piece of wax, Descartes is not suddenly assuming that physical objects exist. Not until the Sixth and last *Meditation* will Descartes overcome his doubt concerning the existence of the physical world. So, the entire discussion of the wax in *Meditation II* is hypothetical: Descartes is considering the question, if physical things did exist, what would they be? Second, although the argument from change and the substance theory are more easily grasped by reference to examples of physical objects such as the wax, Descartes would apply exactly the same reasoning to the purely mental or spiritual thing that he will call the mind or soul. At this point in the *Meditations*, Descartes has not yet established what he later will try to prove—that the mind is a purely mental or immaterial substance. But to avoid a likely, yet radical, misunderstanding of Descartes, it is crucial to realize that for Descartes the argument from change and the substance theory apply to such a nonmaterial "thing" no less than to a material one. Indeed, his use of wax as an example in giving the argument from change is, in one way, misleading: the thing whose existence the *cogito* is supposed to prove is *not* a physical substance but a nonphysical, purely mental one. It is because we (Descartes's readers) are so familiar with physical objects and because Descartes has, in any case, not yet shown that the mind is not physical, that he uses wax as an example. But from the point of view of the doctrine about the mind toward which he is building, it would have been more accurate for him to apply the argument from change directly to the mind. And in the *Synopsis of the following six Meditations* that he prefaced to the *Meditations*, where he is not trying to work toward his doctrines in a step-by-step manner but only summarizing some of his main results, he does just that:

[T]he human mind . . . is a pure substance. For even if all the accidents of the mind change, so that it has different objects of the understanding and different desires and sensations, it does not on that account become a different mind. (CSM II 10, SPW 74, AT VII 14)

The term “accidents” here refers to *properties* of a certain sort—namely, accidental properties. (Accidental properties are those that a thing happens to have but need not have in order to be what it is, e.g., a certain triangle’s property of being blue. Accidental properties contrast with essential properties, which are those that a thing must have in order to be what it is, e.g., the triangle’s property of being three-sided.) Further, Descartes’s phrase “different objects of the understanding and different desires and sensations” here refers to specific, determinate thoughts. So Descartes identifies the accidental properties of a mind with its specific, determinate *cogitationes*. But if a mind’s accidental properties are specific or determinate thoughts, then those properties are also determinate properties. Therefore, Descartes’s argument in the above passage could be briefly summarized as “a human mind is a substance, since even if all its determinate properties change, it is still the same mind.” This is simply the argument from change, now applied directly to the mind.

5. A Reconstruction of the *Cogito* Based on the Substance Theory

We are now in a position to show how the substance theory relates to the *cogito*. We shall do this by giving a step-by-step reconstruction of the *cogito* that uses the substance theory as its very first premise. The first premise of our reconstruction, then, is

- (1) A thing is composed of its properties or characteristics *plus* an underlying substance to which they belong.

To understand this premise, it is important to grasp what the word “thing” means in it. After all, “thing” is the vaguest noun in the English language. Depending on the context, it can be made to stand for virtually anything, including shadows, surfaces, edges, black holes, symphonies, ideas—the list could be prolonged indefinitely. Now in (1), “thing” is being given a specific meaning: it signifies *an entity that could conceivably exist independently of all other entities*. The reason for assigning this meaning to “thing” in (1) is that it is in this sense that the word is understood in the philosophical question that the substance theory claims to answer, namely, “What is a thing?” This question is not about such “things” as shadows, surfaces, borders, or other entities that could not

conceivably exist apart from some other entity. It is a question about what some philosophers call “concrete things”—entities that could at least conceivably exist on their own, apart from any other entities.

What premise (1) says, then, is that *the minimum* that can exist on its own is always a substance and its properties. There could not be a “free-floating” property, that is, a property existing without a substance to which it belongs; nor could there be a substance without any property. So there follows directly from premise 1 (i.e., as a corollary of the substance theory),

- (2) If there is a property or characteristic, then there must be a substance to which it belongs.

This corollary, which rules out “free-floating” properties and is sometimes called the “substance-property principle,” is explicitly put forward by Descartes in several places. For example, in his *Reply* to the fourth set of *Objections*, he says,

[A]ttributes [= properties] . . . must inhere in something if they are to exist; and we call the thing in which they inhere a “substance.” (CSM II 156, AT VII 222)

Likewise, in his *Principles of Philosophy* (a work that presents, in a more didactic manner, the same system of thought as the *Meditations* and also includes much of Descartes’s physics), Descartes says:

[W]e should notice something very well known by the natural light: nothingness possesses no attributes or qualities. It follows that, wherever we find some attributes or qualities, there is necessarily some thing or substance to be found for them to belong to. (Part I, #11 [CSM I 196, SPW 163, AT VIIIA 8])

[N]othingness possesses no attributes, that is to say, no properties or qualities. Thus, if we perceive the presence of some attribute, we can infer that there must also be present an existing thing or substance to which it may be attributed. (Part I, #52 [CSM I 210, SPW 177, AT VIIIA 25])

These quotations call for a short digression. Notice that in them, Descartes tries to deduce (2) from the principle that “nothingness possesses

no properties”—a principle which was a commonplace of medieval philosophy. His argument seems to be that

- (i) Nothingness has no properties.
 - (ii) If there were a property that did not belong to anything, then nothingness would have a property.
-
- . . . (iii) Every property belongs to something.

This, however, is a weak argument. For even if we grant (i), (ii) is not obviously true. Why should the fact that there is some “free-floating” property—one that does not belong to anything—be thought to imply that this property belongs to *nothingness* (i.e., to nonbeing)? Furthermore, (iii) does not imply that the “something” to which a property belongs is a *substance*. Why could it not just belong to a thing, conceived as a bundle or cluster of properties? So it seems that (2) cannot really be established in the way Descartes here suggests. On the other hand, (2) does not *need* to be established in that way, since, as we have already seen, (2) does follow directly from (1). So the failure of the argument from (i) and (ii) to (iii) does no damage to our reconstruction.

Let us return to that reconstruction. Its next step is an assumption that Descartes makes. This is

- (3) A thought is a property.

This assumption is quite explicit in the passage quoted earlier from the *Synopsis* of the *Meditations*, where Descartes characterizes thinking, willing, and perceiving certain things as “accidents” (i.e., accidental properties, as we saw) of the mind. For Descartes, then, a person’s thought of an apple is related to the person in the same way as an apple’s roundness is related to the apple: just as an apple’s roundness is a property of the apple, a person’s thought is a property of the person. This assumption seemed obvious to Descartes; let us reserve comment on it until the next section.

An important consequence of the Assumption and the Corollary (premises 3 and 2, respectively) is

- (4) If there is a thought, then there is a substance to which it belongs.

Descartes himself, in responding to Thomas Hobbes's *Objections to Meditation II*, invokes this principle, deducing it from (2):

[I]t is certain that a thought cannot exist without a thing that is thinking; and in general no act or accident can exist without a substance for it to belong to. (CSM II 124, AT VII 175–6)

The general statement after the semicolon is (2) and is used to support the statement before the semicolon, which is (4). (“Thing” must here mean “substance”; otherwise the first statement would not follow from, or be a particular application of, the second.) The other premise required to deduce (4) from (2), namely (3), is left unstated or “understood” by Descartes.

Next, remember that the premise of the *cogito*, once the apparently question-begging “I” is expurgated from it, becomes simply:

(5) There is a thought.

This much at least, Descartes claims to know indubitably: recall his doctrine that each of us can have absolutely certain knowledge of our own present thoughts, and remember also that (5) is the premise that Bertrand Russell was prepared to “give” to Descartes.

Finally, Descartes moves from (4) and (5) to the conclusion

(6) There is a substance to which this thought belongs: “I.”

This completes the reconstruction of Descartes's argument.

This reconstruction, we suggest, faithfully captures at least one way in which Descartes explains his *cogito*. The least that can be said for it is that no other construal (certainly none that treats the *cogito* as something other than an inference) shows as well how the *cogito* relates to metaphysical principles that Descartes accepted, and that he himself invoked when defending the *cogito* against objections. The reconstruction also shows how Descartes might have defended his *cogito* against the objection that it is either question-begging or invalid. For it shows how he could have thought it possible to prove his existence from the premise “There is a thought,” with the help of certain further premises that he accepted. Anthony Kenny, whose discussion of the *cogito* can be highly recommended, summarizes the argument in essentially the same way as our steps (2)–(6):

When he thinks, he is aware of a thought—no matter, yet, to whom or what the thought belongs [= (5)]. Since, by the light of nature [“by the light of nature” means, for Descartes, “by the light of reason”—i.e., by reason employed in the very best way], he knows what a thought is, he knows that it is an attribute and not a substance [= (3)]. Again, by the light of nature, he knows that every such attribute must belong to a substance [= (2) and (4)]. So he concludes to the existence of the substance of which the thought he perceives is an attribute. This he calls *ego*; or, if you like, he concludes that the “I” in “I am thinking” does refer to a substance and is not just a grammatical convenience. [= (6)]⁵

Kenny does not point out that (2) derives from (1) and is thus a corollary of the substance theory; the logical starting point of his reconstruction is the second step of ours. Aside from this difference, the two reconstructions are the same.

6. Critical Discussion of the Reconstructed *Cogito*

Let us now inquire whether the *cogito*, as we have reconstructed it, is successful. We shall proceed by commenting separately on each of its various steps.

6.1 The Substance Theory and the Argument from Change

The substance theory is not without its problems. The main difficulty with the theory is that substance is unperceivable. Imagine, for example, that you want to see the substance of a block of wood. So, you obtain a carpenter’s plane and plane away a layer of wood. What do you then see? Well, you see a new set of properties—a new size (slightly smaller), shape, and color. It is obvious that planing away still another layer will not get you any closer to seeing the block’s underlying substance. No matter how many layers you plane away, you will see only more properties—until, in the end, all the wood is gone. What this kind of thought-experiment shows is that substance is just not something that can be perceived: nothing would even count as perceiving it. Substance is, as philosophers say, unperceivable *in*

⁵ Anthony Kenny, *Descartes*, p. 60.

principle. For this reason, Empiricist philosophers in the seventeenth and eighteenth centuries became increasingly critical of the substance theory; and many contemporary philosophers reject substance altogether.

In order to reject the substance theory, however, one must refute the argument for it, namely the argument from change. This is not easy to do, and we shall not attempt to do it in this work.⁶ But we can profitably make some remarks about the prospects of refuting the argument.

Premise 1 can be challenged, in the following way. Can we really distinguish between *all* of a thing's determinate properties changing and the thing ceasing to exist? In order for a thing to continue to exist, must it not retain at least some of its determinate properties—if only a single one? And isn't this condition in fact satisfied in the case of Descartes's piece of wax?—for although the wax's observable properties (e.g., its size, shape, color, texture) have all changed, doesn't the wax still retain a certain chemical composition—or perhaps a certain atomic structure—in virtue of which it continues to exist despite the change?

A possible defense against this line of attack can be given by appealing to the following thought-experiment. Suppose that chemists or atomic physicists informed us that when a piece of wax is melted, its color, shape, hardness, and other observable properties are not the only ones that change. Rather, its chemical composition and/or atomic structure are *also* altered, perhaps only very slightly. Would we then feel bound to agree that the piece of wax Descartes describes ceases to exist when we melt it? It may well seem that we would not—at least not merely on the strength of what we have supposed so far, namely, that both the wax's more obvious properties and its “scientific” ones have altered; for if, by cooling and molding the melted stuff, we could easily get it back to its former shape, size, color, texture, and consistency, use it as a candle or a water-repellent, and so forth, then would we not still consider it to be wax? If so, then it seems that what accounts for the wax's continuing to exist when melted is not that it retains “scientific” properties such as its chemical composition or atomic structure. Those could change, as well as the more obvious properties of shape, size, hardness, color, and so on. Also, if the wax, when exposed to heat, *kept* its more obvious properties and only its scientific properties changed, would it not still be wax? So, premise 1 may well seem to withstand the challenge we raised against it.

⁶ There is a fuller critical discussion of the argument from change in Georges Dicker, *Hume's Epistemology and Metaphysics: An Introduction*, pp. 21–31.

Some philosophers, whom we may call “scientific realists,” would not be impressed by this possible response. They would insist that retaining some determinate scientific property is what enables a thing to persist through change. They would support their position by saying that if, for example, a quantity of liquid lost the molecular structure H²O, then it would cease to be water, no matter how much it might still be like water, quench our thirst, and so forth; or, to take another example, that if a piece of yellow shiny metal ceased to have atomic number 69, it would cease to be gold—it would become “fool’s gold.” If this line of thought is correct, then it seems to show that premise (1) of the argument from change is false.

However, we shall not further pursue the question of whether that premise is true or false, because even if scientific realists are right in saying that a thing’s retaining a certain determinate scientific property is a necessary condition for it to persist through change, it would not follow that substance is not needed to give a satisfactory account of what is required for a thing to retain its identity through change. For although scientific realists may be right in saying that a thing’s retaining a certain determinate scientific property is a *necessary* condition for it to persist through change, it cannot be a *sufficient* condition. To see why, consider the following thought-experiment. Suppose that a gold statue occupies a place P1 at a time t, that at time t2 it vanishes from that place, and that at time t3, an exactly similar gold statue begins to occupy a place P2. Further, suppose that it is *not* the case that the statue traveled or was carried from P1 to P2; rather, the gold statue simply ceased to occupy P1 and, later, an exactly similar gold statue began to occupy P2. Further, the statue at P1 has the same atomic number, 69, as the one at P2. Now let us ask: was the statue at P1 numerically identical with the statue at P2? It is important to understand what this question means. It is not asking whether the statue at P1 was *qualitatively* the same as the one at P2. By hypothesis, it was: we just said that the statue at P1 was “exactly like” the one at P2, meaning that it had the same weight, color, shape, size, workmanship, and so forth. Furthermore—and this is important—the statue at P2 had exactly the same atomic structure, atomic number 69, as the statue at P1. Thus, the statue at P1 had exactly the same properties, including the “scientific” ones, as the one at P2. The question, then, is only whether the statue at P1 and the one at P2 were *numerically* the same: was there just one statue which existed at P1 at the earlier time and also at P2 at the later time, or were there two statues, one at the earlier time and another at the later time? (Was the situation like that of one person at two different times, or like

that of two identical twins? Was the statue at P1 the *same one* as the statue at P2, or only the *same as* the one at P2?) The answer, it would seem, is that the statue at P1 and the one at P2 were not numerically identical. The reason is that when the statue at P1 ceased to occupy P1, it ceased to exist, since it did not then travel to P2 (or, we may also imagine, to any other place). Further, when the exactly similar gold statue began to occupy P2 at t3, it began to exist, since it had not traveled *from* P1 (or from any other place). But one thing cannot have two beginnings of existence. So, the statue at P1 and the statue at P2 were qualitatively identical, but numerically distinct (like the twins). If this is right, then it shows that retaining the same scientific properties is not a sufficient condition for a thing to endure through change, since the statue at P1 and the statue at P2 both had atomic number 69.

In light of this point, a proponent of the substance theory might revise the argument from change to go as follows:

- (1R) We can distinguish between (a) all of a thing's determinate properties *except* for scientific property *P* (say, the thing's chemical composition or its atomic structure) changing without the thing's ceasing to exist and (b) a thing ceasing to exist and another thing with the same scientific property *P* beginning to exist.
- (2R) We can distinguish between (a) and (b) only if a thing is composed, in addition to its properties, of a permanent, underlying substance.

. . . A thing is composed, in addition to its properties, of a permanent, underlying substance.

Many contemporary philosophers, however, would reject this version of the argument from change as well, for they would reject premise 2R. Drawing on ideas in John Locke's *An Essay Concerning Human Understanding* (1690), they would say that what is needed for a thing to continue existing through time and change is not some sort of mysterious and unperceivable substance but, rather, that it must meet the requirement of "spatio-temporal continuity."⁷ Roughly speaking, the idea is that a thing continues to exist, or retains its identity through time, only if its career in space and

⁷ See John Locke, *An Essay Concerning Human Understanding*, Book II, chapter 27.

time is not interrupted, but is continuous. In other words, it must *not* be the case that, like the “vanishing statue” in our example, the thing ceases to occupy its place, and that a thing with the same properties begins to occupy a different place (or the same place later), without the thing’s moving from one place to the other (or returning to the same place later).

Although Locke endorses the requirement of spatio-temporal continuity, he argues, in effect, that it too is only a necessary condition for identity through time. Locke argues that a thing must also remain the same *sort* of thing; it cannot, for example, go from being a piece of paper to being a mound of ashes, even if the paper is spatiotemporally continuous with the ashes—even if, say, the paper is ignited and the burning paper is carried from place 1 to place 2, at which point it has become just ashes. This additional requirement yields the now-popular idea that identity through change may be defined, without any appeal to substance, as “spatio-temporal continuity under a sortal.”

There is reason to think, however, that Descartes would not agree that the above considerations refute the argument from change. To see why, we need to look first at a second possible objection to the original version of the argument from change, and then to see how Descartes would answer it. This possible objection is directed against premise (2)—that “we can distinguish between (a) all of a thing’s determinate properties changing without the thing’s ceasing to exist, and (b) a thing’s ceasing to exist only if a thing is composed, in addition to its properties, of a permanent, underlying substance.” The objection is that this premise should be replaced with this one: “we can distinguish between all of a thing’s determinate properties changing without the thing’s ceasing to exist, and (b) a thing’s ceasing to exist only if a thing is composed, in addition to its properties, of at least one determinable property.” The entire argument would then go this way:

(1) We can distinguish between (a) all of a thing’s determinate properties changing without the thing’s ceasing to exist, and (b) a thing’s ceasing to exist.

(2R*) We can distinguish between (a) and (b) only if a thing is composed, in addition to its properties, of at least one determinable property.

∴ A thing is composed, in addition to its determinate properties, of at least one determinable property.

Now this possible objection will not do, because (2R*) is false. For it does not seem that merely retaining some determinable property—for example some shape or other, or some size or other, or some color or other, and so forth—is a sufficient condition for a thing to retain its identity through time though all of its determinate properties have changed, even if the requirement of spatio-temporal continuity is also satisfied. Suppose, for example, that your car were put into a powerful compactor and reduced to a six-by-six foot solid cube of metal: that cube would still have a shape and size and a color, and it would be spatio-temporally continuous with what had been your car, but surely your car would no longer exist.

Nevertheless, as we shall see more fully in chapter 6, there is good evidence that Descartes would not only have *accepted* the reformulated argument that uses 2R*, but would also have seen it as actually *establishing* his version of the substance theory. For ultimately he understands “thing” in such a way that a physical thing could lose of all of its determinate properties and yet remain the same thing. It could do so by virtue of retaining a single determinable property, namely, the property of shape-and-size, which he calls “extension.” According to the view that Descartes seems ultimately committed to, an ordinary object like a piece of wax or a rock or a car is not a “thing” in its own right, that is, an entity that could conceivably exist apart from all other entities. Rather, it is only a cluster of properties belonging to a single, all-encompassing physical thing. This all-encompassing physical thing is nothing less than the entire physical world—the entire physical universe. The piece of wax discussed in *Meditation II* must then be seen as only a miniature model that Descartes uses to represent the entire physical world. And the point of the argument from change, as he there presents it, is that so long as the physical world retains the one determinable property of *extension*, it continues to exist as one and the same physical world. On this view, furthermore, physical substance is *identical* with the determinable property of extension. More carefully stated, it is identical with the one instance or occurrence of (the determinable property of) extension that exists. Since, on this view, the determinable property of extension is identical with (physical) substance, the reformulated argument from change still establishes the substance theory. Furthermore, premise 1 of the argument now seems to be on solid ground: it is no longer even relevant to object, as we did against the original version of the argument, that in order to persist through change, a thing must retain at least one determinate “scientific”

property such its chemical composition or its atomic structure. For all that is needed, instead, is that it retain the determinable property of extension (three-dimensionality), which it can do even if all of its determinate properties, including both its readily perceivable ones and its scientific ones, have altered. Finally, substance must no longer be thought of as a mysterious something underlying all of a thing's properties. Instead, it is just the thing's most basic, defining property: extension in the case of matter, and thought in the case of mind. So, the substance theory is no longer vulnerable to the standard objection that substance is in principle unperceivable.

As will be shown in chapter 6, this way of conceiving the matter is arguably the most accurate interpretation of Descartes's view of substance. But philosophically, there is a high price to be paid for it. This is that one must give up the view that ordinary objects are things in their own right, in favor of the view that they are just clusters of properties of the one all-encompassing physical substance. Not many philosophers would be willing to go that far away from common sense. So, unless philosophers who reject such a one-substance view of the physical world want to accept the theory of an unperceivable substance underlying a thing's properties (which most do not), they have to find some other way of dealing with the argument from change, by giving some account of identity through change that does not rely on substance. As we have seen, some contemporary philosophers have in effect tried to do this, by appealing to the concept of spatio-temporal continuity under a sortal.⁸

6.2 The Corollary

The Corollary of the Substance theory expresses a certain view about the status of properties: it says that a property cannot exist "on its own," but rather can exist only in a substance. Although, at first, this view may seem unproblematic, it is not the only possible view about the status of properties. In fact, it is one of three competing views. To understand these views, notice first that what is distinctive about a property is that it can be had by several different things. The property of white or whiteness, for example, can be had by a sheet of paper, a

⁸ For a defense of this kind of view see Dicker, *Hume's Epistemology and Metaphysics*, pp. 24–31.

piece of chalk, a snowflake, and indefinitely many other things. By contrast, the paper, chalk or snowflake cannot be similarly “had” by different things (though they might be used, or even legally owned, by several different persons). Thus, what is special about properties is just this: unlike particular, individual things, a property can be shared, or had in common, by many different things. This raises a puzzling question: How is it possible for *many* different things, which may be spatially far removed from each other, to share *one* property (e.g., for a snowflake, a piece of paper, and a piece of chalk all to be white)? Each of the three views just mentioned attempts to answer this question, and the problem of deciding between the three views is called “the problem of universals.” (In presenting this problem, the noun “universal” means the same as the word “property.”) The problem of universals is a hard one, which we shall not try to solve. We shall only summarize the three competing solutions, and point out some problems associated with each of them.

1. According to *Platonic Realism*, universals (properties) exist independently of particular things. To say that two or more things share a common property is to say that they stand in a special relationship—usually called “exemplification”—to the universal. Thus the snowflake, paper, and chalk each “exemplify” whiteness, which exists independently of them and of other particular white things. According to Platonic Realism, then, properties *are* “free-floating”; and the corollary of the substance theory (as well as the theory itself) is false. The term “Platonic Realism” derives from the name of this view’s originator and principal exponent, the Greek philosopher Plato (427–347 b.c.).
2. According to *Moderate Realism*, universals exist only in particular things. To say that two or more things have a common property is to say that the property somehow exists *in* those things. Whiteness, for example, does not exist on its own apart from white objects; it exists in the snowflake, the paper, the chalk, and other white things. Moderate Realism is no doubt closer to common sense than Platonic Realism. It is also the view implicit in the substance theory and reflected in its corollary, according to which properties cannot be “free-floating” but must exist in a substance. Historically, Moderate Realism derives mainly from Aristotle (384–322 b.c.), who was Plato’s student and who broke from his master precisely on the fundamental question of the status of universals.

3. According to *Nominalism*, only particular things exist. Universals do not really exist, either independently of particular things or in those things. To say that two or more things have a common property (e.g., are white) is only to say that the things *resemble* each other in a certain way, and that we therefore apply the same *name* to them (“Nominalism” derives from “name”). Thus, our ontology need not include universals; particulars that resemble each other in various ways are all that really exists. Some classical defenders of Nominalism are the English philosophers Hobbes (1588–1679), Berkeley (1685–1753), and Hume (1711–1776).

Each of the above theories faces difficulties. The most obvious difficulty for Platonic Realism is simply that independently existing universals are very mysterious entities; for, unlike particular objects, they cannot be located in space. For example, whiteness itself—unlike the snowflake, the paper, and the chalk—does not have a spatial location and does not occupy a volume of space; for if it did, how would it differ from particular white things? So with respect to independently existing universals, questions such as “Where is it?” or “How large is it?” simply lack answers.

A less obvious difficulty, which Plato himself anticipated (in the *Parmenides*, one of Plato’s many *Dialogues*, which contain the bulk of his philosophy) is known as the problem of the “Third Man.” It can be put this way. Suppose we ask, Is the universal, whiteness, itself white? If we say that whiteness is white, then in addition to particular white things such as the snowflake, the paper, and the chalk, we have another white object, namely, whiteness. But then, the same question arises about the particulars plus the universal as arose about the particulars alone, namely, What makes it true that these things (including whiteness itself) are all *white*? If we answer that they all exemplify yet another universal, say whiteness₁, then the very same problem recurs with respect to whiteness₁: Is it white? If the answer is yes, then for the same reason we are forced to introduce yet another universal, whiteness₂, and so on *ad infinitum*. On the other hand, if we try to avoid this “infinite regress” (as philosophers call such a series) of whitenesses by saying that whiteness itself is *not* white, then it is extremely hard to see how particular white things can truly be said to “exemplify” whiteness or how the theory provides any answer to the question, What makes it true that white things are all white?

Faced with such difficulties, it is very natural to give up Platonic Realism in favor of Moderate Realism (as did Aristotle, who was thoroughly familiar with the obscurities and problems involved in postulating independently existing universals). But Moderate Realism also faces difficulties. One such difficulty is: How can *one* thing, such as whiteness, be in several different things? Here you may object that whiteness isn't a "thing" but a quality, so that the difficulty is not a genuine one. But saying that whiteness is a quality, rather than a thing, does not eliminate the basic difficulty; for a quality is at least something that exists—an existent—and it is puzzling how any single existent can be in several different things. Although at first it seems just common sense to say that the color white is literally in the snowflake, the paper, and the chalk, the minute we think about the matter seriously, we are bound to find it perplexing.

There is another difficulty for Moderate Realism. Suppose that all particular white things were destroyed. Would whiteness itself cease to be? It seems not; rather, it seems that whiteness would continue to exist no matter what happened to particular white things like snowflakes and bits of chalk. But if so, then we seem forced back to Plato's view that whiteness exists independently of white things.

The difficulties faced by both Platonic and Moderate Realism may well make Nominalism attractive. For according to Nominalism, universals do not really exist at all; so we need not worry about answering the puzzling questions just raised regarding them. Unfortunately, however, Nominalism faces its own difficulties. One problem for Nominalism is that *any* two or more things resemble each other in *some* way—if only in that they all exist, or are in space and/or time. Therefore, in order to explain how a piece of chalk and a snowflake can both be white, it is not sufficient merely to say that the snowflake and the chalk resemble each other. We must say, instead, that the snowflake and the chalk resemble each other in respect of—. But the only way to fill in this blank, it seems, is by referring to a *property*: "in respect of being white." But then, we have reintroduced the universal that it was the purpose of Nominalism to avoid.

Another difficulty for Nominalism, as Bertrand Russell pointed out, is that even if we succeeded in avoiding commitment to universals in favor of resembling particulars, it seems that we would still be committed to the existence of at least one universal; namely, resemblance itself. But if we allow even one universal into our ontology, then the puzzling questions about universals arise with respect to *that* universal (the camel's

nose is let into the tent, so to speak), and so the advantages of Nominalism are lost.

It can be seen, then, that Platonic Realism, Moderate Realism, and Nominalism each face certain difficulties. Yet, it would seem that one of the three views must be correct; for they seem to exhaust the possibilities. To solve the problem of universals, one would have to show that one of the views can be formulated in such a way as to meet the difficulties it faces.⁹

6.3 The Assumption That Thoughts Are Properties

Given an ontology according to which whatever exists is either a substance or a property, it seems natural to classify thoughts as properties. But this naturalness is somewhat deceptive. For it is not easy to say why a thought could not be a substance rather than a property. Of course, if one assumes Descartes's dualistic view of the self, according to which the self is a mental substance that could conceivably exist apart from anything physical, then it becomes possible to argue that since (1) this substance must have *some* properties (by the substance theory), (2) these properties must be mental rather than physical ones (by the theory of dualism), and (3) the only mental items available to fill this role are thoughts, it follows that thoughts are properties. But this line of argument is not available when reconstructing the *cogito*; for Descartes, who at that point is still trying to prove his existence, has not yet established—and is indeed only starting to build toward—his dualism. So some other argument is needed to show that thoughts are properties.

Perhaps the following argument (which is not Descartes's) will do. A substance is something that can acquire and lose properties while remaining the same substance, that is, without losing its identity. But could a *thought* acquire and lose properties *while remaining the same thought*—without losing its identity? It seems not: if, for example, I think of a ship with one chimney and then of an exactly similar ship except that it has two chimneys (or if I think of \$1,000.00 and then think of \$1,000.01), am I not thinking a different thought? If so, then it would seem that thoughts cannot be substances and must, therefore, be properties.

⁹ There is an excellent discussion of the issue, featuring a defense of nominalism, in chapter 1 of H. H. Price, *Thinking and Experience*.

Descartes, in any case, took it as obvious that thoughts are properties, rather than substances. We shall not comment further on this assumption, except only to note that it *is* an assumption Descartes makes. Thus, the reconstructed *cogito* assumes not only the truth of the substance theory but also the correctness of a particular application of it. It has become apparent, then, that despite Descartes's wish to doubt everything that is not absolutely certain, even in expounding the *cogito* he makes assumptions that might be questioned. Concerning such assumptions, it has been well said by Richard Schacht that:

Descartes did not question these assumptions, not because he was deliberately trying to cheat, but rather because they were such fundamental assumptions of medieval thought that it simply did not occur to him that they *were* assumptions standing in need of justification, and perhaps untenable ones. Descartes broke rather significantly with medieval philosophical thought in his program and proposed method; but he still took for granted a number of very basic axioms of medieval reasoning. (One encounters more of this sort of thing in his proofs of the existence of God.) It remained for subsequent philosophers to notice that the method he proposed, and the standards of knowledge he set, would undercut not only the claim of many of our commonsense opinions to be genuine knowledge, but also many of the basic axioms or assumptions which he still retains. This does not establish that they are actually *wrong*; it only shows that his proposed method and standard had consequences he did not foresee.¹⁰

6.4 The Inference to “I Exist”

Suppose that we accept all the premises from which, according to our reconstruction of Descartes's reasoning, he derives his existence. Does the argument then succeed in proving Descartes's existence? No, it does not. For nothing in the argument entitles Descartes to call the substance whose existence is derived in the argument's last step, “I.” *At best* the argument only proves the existence of a thinking substance; it does not

¹⁰ Richard Schacht, *Classical Modern Philosophers*, p. 19.

prove that this substance is myself. The first-person pronoun, “I,” so far as the logic of the argument goes, is a completely gratuitous addition to the conclusion drawn in step (6). The criticism that we are here making has been incisively put by Anthony Kenny. He first notes, in the same spirit but rather more pointedly than Schacht, that:

[T]he principle that where there are attributes there must be a substance does not seem as unquestionable since the writings of Berkeley and Hume as it did to Descartes. Too often, when Descartes tells us that something is taught by the natural light in our souls, he produces a doctrine taught by the Jesuits at La Flèche [La Flèche was the school Descartes attended as a youth].

Kenny then unleashes the telling objection to Descartes’s conclusion:

But even if we accept the principle, there seems some doubt whether the conclusion it licenses is in fact “*sum*.” [*Sum* is Latin for “I am.”] Is not Descartes rash in christening the substance in which the doubts of the *Meditations* inhere “*ego*”? To be sure, he explains that he is not yet committing himself to any doctrine about the nature of the *ego*; not until the *Sixth Meditation*, for instance, will he prove that it is incorporeal. But what “I” refers to must at least be distinct from what “you” refers to; otherwise the argument might as well run “*cogitatur, ergo es*” [Latin for “thinking is occurring, therefore you exist”] as “*cogito ergo sum*.” Has Descartes any right to make such an assumption about the substance in which these thoughts inhere? In 1641 Hyperaspistes wrote: “You do not know whether it is you yourself who think, or whether the world-soul in you thinks, as the Platonists believe.” To this pertinent criticism Descartes had no real reply.¹¹

To elaborate (but also to make one small critical observation) on Kenny’s criticism, the reconstruction of the *cogito* from the substance theory attempts to derive the conclusion “I exist” from the premise “there is a thought,” taken in conjunction with several metaphysical principles, (1) through (4), all of which are stated *impersonally*, that is, without any explicit or implicit reference to the self or “I.” Kenny’s key point is that from

¹¹ Kenny, *Descartes*, pp. 61–62.

such premises, the conclusion “I exist” does not logically follow; at best, what follows is that there exists some thinking substance to which said thought belongs. He illustrates this point by asserting that the “I” of the conclusion might just as well be “you,” so that Descartes is “rash in christening the substance in which the doubts of the *Meditations* inhere ‘ego.’” He also quotes Hyperaspistes’s remark that “you do not know whether is it you yourself who think, or whether the world-soul in you thinks.” It would seem, however, that the case of the world-soul is not a good illustration of Kenny’s point: he should have stuck with his first illustration only; for the statement that “the world-soul is thinking in me” does entail that I exist (due to the “in me”). Still, this does not affect the basic point that “I exist” doesn’t follow from the premises of the reconstruction, anymore that “you exist” or “he/she exists.” From those premises, nothing at all follows about the identity of the thinking substance in which the thinking occurs.

7. A Defense of the Unreconstructed *Cogito*

In light of the failure of the reconstructed *cogito*, let us reexamine the original charge that the basic, unreconstructed argument, “I am thinking, therefore I exist,” is question-begging. The reason given above for this claim went as follows:

What entitles Descartes to use the first-person pronoun “I” in the premise of his proof? Descartes’s use of this pronoun presupposes the very thing he is supposedly proving, namely, that he exists. Thus the *cogito* appears to be what is called a *question-begging argument*, that is, an argument that takes for granted or assumes the very conclusion that it is supposed to prove.

The apparent question-beggingness of Descartes’s proof stems from the special semantics of the pronoun “I.” This pronoun is what we may call a *uniquely referring term*. To see what is meant by this, compare “I” with “you,” “he,” or “she.” The referents of “you,” “he,” or “she” may be the same when uttered by different persons; for example, you and I could both refer to Hillary Clinton by means of the pronoun “she” (if we were both talking about her) or by means of the pronoun “you” (if we were both talking to her). But the referent of “I” must be different when uttered by different persons; you and I cannot both use “I” to refer to me; Hillary Clinton and

Michelle Obama cannot both use “I” to refer to Hillary, and so on. Furthermore, the referents of “you,” “he,” and “she,” when uttered by the same person, can obviously vary, and can be determined only when some contextual factor other than the mere utterance of those pronouns serves to clarify who is being referred to. By contrast, no such clarification is needed to determine who “I” refers to: its referent is always the very person who utters it. Even if two different actors both utter “I” while playing Hamlet, their uses of “I” refer to one and the same character, namely Hamlet. Simply put, the pronoun “I” has only one use—it refers to or calls attention to oneself. These points, or more briefly the point that “I” is a uniquely referring term, seem to be all that lies behind the charge that in saying “I am thinking,” Descartes is already referring to himself, and so already assuming or taking it for granted that he exists.

It is arguable, however, that this objection is too hard on Descartes. First, note that the definition of a question-begging argument as one “that takes for granted or assumes the very conclusion that it is supposed to prove,” although it is the standard definition, is quite loose. An argument is only a set of *statements* such that some of those statements are given as reasons or grounds for another statement in the set. (The context, or the presence of such words as “therefore,” “hence,” “consequently,” “since,” “because,” “for the reason that,” and so forth, indicate that the statements are being so treated.) As such, an *argument* cannot literally “take for granted” or “assume” anything: only an *arguer*—a person who produces an argument—can do that. It seems, then, that a more careful definition of a question-begging argument is needed. We need to recognize that the notion of begging the question is essentially an epistemological one.¹² We propose the following definition:

a question-begging argument = *df* an argument such that it is impossible to know the premise(s) of the argument without already knowing its conclusion.

In light of this definition, and of our points about the special semantics of “I,” we can give a plausible defense of the *cogito* against the charge that it begs the question, for it seems possible for a person to know “I am thinking” without already knowing “I exist.” This could come about if the following two conditions were both satisfied:

¹² I owe this insight to Gary Iseminger.

(1) A person might know that “I” always refers to the self without grasping that the self exists. For there are uniquely referring terms, notably proper names, that refer to nonexistent things, e.g., “Superman,” “Batwoman,” “Hamlet” (when used to refer to Shakespeare’s fictional prince), “Shangri-La,” and the like, as well as pronouns like “he,” “she,” “it,” and even “you,” that refer to nonexistent things, as when they are used to refer to Superman, Batwoman, Hamlet, Shangri-La, or to one’s imaginary muse. *There is nothing about the pronoun “I” that demonstrates that, unlike a proper name or a singular pronoun, it can only refer to something that actually exists.* The pronoun “I” does not magically create its own referent. But in that case, just as Descartes insists, it is the actual thinking in the *cogito* that proves one’s own existence: the thinking certainly and really exists—it is manifestly occurring now, but it is *my* thinking, so I must exist. If someone complains that the thinking may not be mine, Descartes has an intuitively compelling answer: “the fact that it is I who am doubting and understanding and willing is so evident that I see no way of making it any clearer” (CSM II 19, AT VII 29, SPW 83). His point here could be put by saying that the “mineness” or first-personal nature of my own thinking—the fact that it is my thinking, not anyone else’s, or no one’s—is given, that it is a datum.¹³

This point does not mean that I might not know that I exist—that is a separate issue, on which we need not take a position. Rather, it only means that I can know how the pronoun “I” functions, and use that pronoun, without *thereby* knowing that I exist. Of course, I cannot assert that I exist unless I do in fact exist, but neither can I assert anything else unless I exist: this is just a general point about the necessary conditions of my making any assertion whatsoever that has no special relevance to the *cogito*; indeed it is just an illustration of Gassendi’s remark that “you could have made the same inference [that you exist] from any one of your . . . actions, since it is known by the natural light that whatever acts exists” (CSM II 180, M 68, SPW 126, AT VII 259)—a remark which, as we saw earlier, does not count against the *cogito* since it ignores the special certainty of “I am thinking” and of other *cogitatio*-statements.

¹³ There is a difficult discussion that may present a challenge to this claim in Bernard Williams, *Descartes: The Project of Pure Enquiry*, pp. 95–101.

(2) A person might know “I am thinking” without ever having had the thought that this entails “I exist,” or ever having had the thought that “if I am thinking, then I exist.” Prior to drawing the conclusion “I exist” from the premise “I am thinking,” it might never have occurred to her that her thinking entailed her existence. Of course, after going through the *cogito* just once, a person might never again have the thought “I am thinking” without also realizing that having this thought necessitates that she exists. But there is no reason why that should have bothered Descartes; quite the contrary, it harmonizes well with his view that the *cogito* can serve as a stable foundation for other knowledge.

It might now be objected, however, that “I am thinking” begs the question because the use of “I” *covertly asserts* that I exist. Now it seems that the point already made about names and pronouns that refer to nonexisting things is enough to refute this objection too. But the objection also runs into another problem, because it seems reasonable to say that in *any* valid argument, the premises covertly assert the conclusion. Take for example the classic syllogism:

All humans are mortal
Socrates is human

. . . Socrates is mortal.

Then if we complain that “I am thinking” covertly asserts “I exist,” should we not also complain that “all humans are mortal and Socrates is human” covertly assert “Socrates is mortal”? Or take the *modus ponens* argument:

If it rains, then the ground gets wet
It rains

. . . The ground gets wet.

Again, if we complain that “I am thinking” covertly asserts “I exist,” then should we not also complain that the premises of this argument covertly assert its conclusion? But if the answers to these questions are yes, then we seem to be committed to the repugnant conclusion that *all* logically valid arguments are question-begging.

We can say more on behalf of the *cogito*. Even if you are not convinced by the above considerations that it eludes the charge of being question-begging, you should not conclude that the entailment from “I am thinking” to “I exist” is unimportant. On the contrary, there are at least two different reasons why this entailment would remain important even if the *cogito* were question-begging. To appreciate the first reason, try to put yourself in Descartes’s frame of mind at the beginning of *Meditation II*. There Descartes was in a state of radical, disorienting uncertainty, as was conveyed by his image of the man caught in a deep whirlpool who can neither touch the bottom nor swim to the surface; for he was uncertain of the very existence of the entire physical world, including even his own body. Now what the entailment of “I exist” by “I am thinking” shows is that even if all my beliefs about the material world, including even those about my own embodiment, are uncertain, my existence remains certain; for it still remains certain that I am thinking. But just from this one very meager certainty, it already follows that I exist. Thus, even in the midst of the most extreme uncertainty, one can become perfectly certain of one’s own existence. So even if one thinks that the entailment of “I exist” by “I am thinking” does not, strictly speaking, amount to a proof that one exists due to the issue of question-begging, it still shows something important: even in the face of the extreme, disorienting doubt generated by the arguments of *Meditation I*, one’s own existence remains unshakably certain. Or, to put it another way, even doubting the existence of my own body is not tantamount to doubting *my* existence.

The second reason why the entailment is important is this. If “I am thinking” entails “I exist” but does not entail “I have a body,” then “I exist” does not entail “I have a body.” Now this suggests (though it does not by itself prove) Descartes’s view that a person could exist without a body, as a mere thinking thing or disembodied mind. Thus the *cogito*, as Descartes intimates in many places, serves as a springboard to his mind-body dualism—the famous doctrine about the self to which we shall turn our attention in section 9.

8. Does the Unreconstructed *Cogito* Require an Additional Premise?

Some of Descartes’s contemporaries criticized the *cogito* from an angle that we have not yet discussed. They claimed that the argument

I am thinking

\therefore I exist

is simply invalid as it stands. They suggested that Descartes should have added the premise “whatever thinks exists,” or some similar premise, to make the argument valid. Their point can be put by noting that, from a strictly logical point of view, the *cogito* is invalid because it has the logical form

p

\therefore q

which, of course, is invalid.¹⁴ Now the obvious way to fix the argument (assuming that it needs fixing) would be to add the premise “if I am thinking, then I exist,” thereby turning the argument into one with the form

If p, then q

p

\therefore q

which, of course, is valid (*Modus Ponens*); or to add some similar premise, such as “whatever thinks exists,” that would make the argument valid.

¹⁴ In modern quantificational logic, which did not exist in Descartes’s day, the *cogito* could be symbolized as

Ba

$\therefore (\exists x)[Bx. (x = a)]$

But, as Hintikka points out, in systems of logic in which this pattern of inference is valid (or in which the principle “ $\{Ba \supset (\exists x)[Tx. (x = a)]\}$ ” is provable), it is assumed that “the term which replaces $a \dots$ must not be empty. It turns out, therefore, that we in fact decided that the sentence “I exist” is true when we decided that the sentence “I think” is of the form $B(a)$.” Jakko Hintikka, “*Cogito, Ergo Sum: Inference or Performance?*” p. 113. Thus, if the *cogito* is symbolized in this way, then it is valid but question-begging.

So long as we were supposing that the “I” in “I am thinking” begs the question, the issue of whether the *cogito* needs another premise was moot. When an argument is question-begging, it cannot possibly require another premise to make it valid, whether we use the standard definition of begging the question in terms of the premises taking the conclusion for granted, or our proposed, improved definition in terms of the impossibility of knowing the premises without already knowing the conclusion. For part of any explanation of *why* an argument is question-begging is that its premises entail its conclusion: only that fact could explain why the premises take the conclusion for granted, or why it is *impossible* to know the premises without knowing the conclusion. But if the premises of an argument entail its conclusion—if the argument is logically valid as it stands—then it obviously cannot require another premise to make it valid.

However, now that we are no longer supposing that the *cogito* begs the question, we can sensibly raise the question of whether it requires an additional premise. Descartes’s own reply to this question was, it seems, not wholly consistent. On the one hand, in the *Principles of Philosophy*, Part I, section 10, he seems to concede that the *cogito* presupposes an additional premise:

And when I said that the proposition *I am thinking, therefore I exist* is the first and most certain to occur to anyone who philosophizes in an orderly way, I did not in saying that deny that one must first know what thought, existence, and certainty are, and that it is impossible that what thinks should not exist, and so forth. But because these are very simple notions, and ones which on their own provide us with no knowledge of anything that exists, I did not think they needed to be listed. (CSM I 196, SPW 163. AT VIIIA 8)

Here Descartes seems to admit that the *cogito* requires, as an additional premise, the principle that “it is impossible that what thinks should not exist.” But on the other hand, in a number of places Descartes insists that the *cogito* is completely compelling without the addition of any additional premise. Thus in replying to the second set of *Objections*, he says:

And when we become aware that we are thinking things, this is a primary notion that is not derived by means of any syllogism. When someone says, ‘I am thinking, therefore I am, or exist,’ he does not deduce existence from thought by a syllogism,

but recognizes it as something self-evident by a simple intuition of the mind. This is clear from the fact that if he were deducing it by means of a syllogism, he would have to have had previous knowledge of the major premise ‘Everything which thinks is, or exists’; yet in fact he learns it from experiencing in his own case that it is impossible that he should think without existing. (CSM II 100, M 68, SPW 127, AT VII 140)

How are we to reconcile these two passages? We may interpret Descartes as saying, in the second one, that “I exist” is known by an immediate (one-step) inference from “I am thinking.” He adds that one only becomes aware of an additional premise as a result of first grasping the necessary connection between thinking and existing in one’s own case. The added premise is only recognized in retrospect; it is not needed to grasp what Descartes sees as the blinding force of the *cogito*. In other words, the additional, general premise “everything which thinks is, or exists” is grasped only as a by-product of grasping the necessity of the move from “I am thinking” to “I exist” in one’s own individual case. Descartes would presumably say the same thing about the conditional “if I am thinking, then I exist”: it is not needed in the *cogito*, but instead is known only as a result of grasping in one’s own case that one must exist since one is thinking. While this reading of the passage may not render it perfectly consistent with the passage from the *Principles*, it seems to harmonize best with Descartes’s repeatedly-stated claim that “I am thinking, therefore I exist” is his very first piece of knowledge.¹⁵

Note that there is certainly a reason why it is important for Descartes to insist that the *cogito* does not need an additional premise. This is that it seems possible that an evil deceiver could fool him about any such additional premise as “whatever thinks exists” or “if I am thinking, then I exist.” The *cogito* needs to be, so to speak, “deceiver-proof” (using the term “proof” as in “waterproof”), but it cannot be deceiver-proof if it requires a premise other than “I am thinking.”

It remains true, of course, that the *cogito* is formally valid only if the additional premise is added. But Descartes’s position is that the additional premise need not be already known in order for one to grasp the necessity of the connection between one’s own thought

¹⁵ This claim is still not quite correct, since Descartes admits that one must first know “what thought, existence, and certainty are.” But perhaps, as the passage from the *Principles* suggests, his real point is that the *cogito* yields his first knowledge about what *exists*.

and one's existence. That connection is grasped immediately, and the knowledge of the additional premise depends on that grasp, rather than the other way around. Here, so to speak, epistemology is prior to (formal) logic.¹⁶

Before drawing our discussion of the *cogito* to a close, let us consider a final question about it: Is there a way to integrate our defense of the unreconstructed *cogito* into the argument that is supposed to show that I am a *substance*? Indeed there is, and it can be summarized as follows:

1. I am thinking.
2. If I am thinking, then I exist [needed for the inference to (3) to be formally valid, but not needed to know that (3) on the basis of (1)].
3. I exist. [from (1) and (2), though (2) is known only in retrospect]
4. If I am thinking, then I have a thought (call it T).
5. I have a thought T. [from (1) and (4)]
6. If I have a thought T, then there is a thought T.
7. There is a thought T. [from (5) & (6)]
8. A thought is a property. (Descartes's assumption)
9. If there is a thought T, then T is a property. [from (8)]
10. If x is a property, then there is a substance to which x belongs. (corollary of the substance theory)
11. If there is a thought T, then there is a substance to which T belongs. [from (9) & (10)]
12. There is a substance to which T belongs. [from (7) & (11)]
13. If I have a thought T and there is a substance to which T belongs, then I am the substance to which T belongs.
14. I am the substance to which T belongs. [from (5), (12), and (13)]
15. If I am the substance to which T belongs, then I am a thinking substance.
16. I am a thinking substance. [from (14) & (15)]

Here the first three steps constitute the (unreconstructed) *cogito*, while the remaining steps draw on ideas contained in the reconstructed one. The only premise that goes beyond those ideas is (13), and that premise seems very plausible.

¹⁶ Compare with the excellent discussion of this point in Gary Hatfield, *Descartes and the Meditations*, pp. 106–15. See also the fine but somewhat different discussion in Williams, *Descartes: The Project of Pure Enquiry*, pp. 88–92.

9. Descartes's Conception of the Self

So far, we have examined only the three opening paragraphs of *Meditation II*, in which Descartes advances his famous proof that he exists. We now turn to the rest of *Meditation II*, in which Descartes introduces and clarifies his conception of the self. Let's begin with a summary of the main steps Descartes takes in the rest of the *Meditation* (the numbers continue the sequence begun at the start of this chapter):

- 5. I now know that I am, but not yet what I am.**
- 6. I shall therefore review my former beliefs about myself to see if any of them are certain and indubitable. I believed that I was**
 - (a) a man;**
 - (b) a being with a face, hands, arms, etc.,—i.e., with a body; and**
 - (c) a being who was nourished, who moved about, and who engaged in sense perception and thinking—actions that I assigned to the soul.**
- 7. Only thinking indubitably belongs to me; for everything else I've just mentioned depends on my body, which may not exist at all (since there may be an evil deceiver).**
- 8. I am only a thing that thinks (doubts, understands, affirms, denies, wills, refuses, imagines, feels).**

Satisfied that he has proved his existence, Descartes turns in step 5 to the question “*What am I?*” As he puts it, “But I do not yet have a sufficient understanding of what this ‘I’ is, that now necessarily exists” (CSM II 17, SPW 81, AT VII 25). His treatment of this question takes up the rest of *Meditation II* and a substantial portion of *Meditation VI*, as well.

In order to make sure that he will accept no beliefs about himself that are not absolutely indubitable, Descartes proposes, in step 6, to apply the same method to beliefs specifically about himself that he applied more generally in *Meditation I*—the method of doubt. This requires him to examine his former beliefs about himself, in order to see which, if any, of them can withstand the test of doubt. Thus, he says:

I will therefore go back and meditate on what I originally believed myself to be, before I embarked on this present train of thought. I will then subtract anything capable of being weakened, even minimally, by the arguments now introduced, so that what is left

at the end may be exactly and only what is certain and unshakeable. (CSM II 17, SPW 81, AT VII 25)

Accordingly, Descartes now reviews his former beliefs about himself. Some of the points he makes about those beliefs require explanation. Regarding 6(a), he says:

But what is a man? Shall I say ‘a rational animal?’ No; for then I should have to inquire what an animal is, what rationality is, and in this way one question would lead me down the slope to other harder ones, and I do not now have the time to waste on subtleties of this kind. (CSM II 17, SPW 81, AT VII 25)

This is an ironical remark, in which Descartes is expressing his contempt for the Aristotelian method of definition by genus and species—which was standard at Descartes’s time—according to which “man” is defined, famously, as “a rational [species] animal [genus].” Descartes believed that this method of definition only leads to further questions about how to define the genus and the species and that there is a better method, involving “clear and distinct ideas,” for grasping the meaning of important notions. We shall see Descartes putting this method to work in a moment.

Having expressed his dissatisfaction with the prevailing method of definition, Descartes turns his attention in 6(b) to what he spontaneously believed about himself, that is, to beliefs he had acquired quite apart from supposedly learned definitions of “man,” such as the belief that he had a face, limbs, and so on. But in 6(c) he returns to beliefs that stem from Aristotelian conceptions prevailing at his time. To us, it sounds very strange to say that eating, moving around, and sense perception (here conceived not as a type of *cogitatio* but as an activity requiring sense-organs and therefore dependent on the body) are activities of the soul. But this description would not have puzzled Descartes’s seventeenth-century readers, because it is based on the Aristotelian conception of the soul with which they were familiar.

According to Aristotle, the soul is what makes a thing alive; it is the principle of life. So plants and nonhuman animals, as well as humans, have souls. However, there is a difference of degree between the souls of plants, nonhuman animals, and humans; for a soul can have several different parts or faculties, and the “lower-grade” souls lack some of these. Specifically, the souls of plants have only the “nutritive” faculty—the

part of the soul that allows a thing to take in nourishment. The souls of animals, in addition to the nutritive faculty, also possess the sensory, appetitive, and locomotive (motion-originating) faculties, which allow the animal to detect, desire, and move toward food or nourishment. Finally, human souls, in addition to having all of the faculties just mentioned, also have the faculty of rational thought (hence the definition of a human being as a rational animal). In light of this theory of the soul, which Aristotle expounded in a work titled *De Anima* (Latin for *On the Soul*), we can understand Descartes's talk of assigning eating, moving, and sense perception to the soul: these are the activities of the nutritive, locomotive, and sensory faculties of the human soul, respectively.

In step (7), Descartes gives the fundamental reason why he can accept as certain almost none of the beliefs that he has just reviewed. The hypothesis that there may be an evil deceiver, who fools him about the existence of the entire physical world, is still in force. As he puts it,

But what shall I now say that I am, when I am supposing that there is some supremely powerful and . . . malicious deceiver, who is deliberately trying to trick me in every way he can? (CSM II 18, SPW 81, AT VII 26)

The deceiver hypothesis shows that no belief that implies that Descartes has a body can be certain; thus, it shows at one stroke that none of the beliefs listed in 6(a)-6(c) can be certain—with one exception. The exception is Descartes's belief that he is *thinking*. On the ground that there may be a deceiver who fools him about the existence of the whole physical universe, Descartes can doubt that he possesses a body and so that he takes in food or moves around or has sense-perceptions (in the sense involving physical stimuli and sense organs). But the deceiver hypothesis cannot shake Descartes's belief that he is thinking. For, as we saw in section 2 above, while it would be possible for a very powerful deceiver to fool Descartes into thinking that he had a body even if he didn't have one, it would be impossible for any deceiver, no matter how powerful, to fool Descartes into thinking that he was thinking when he wasn't thinking—for thinking that one is thinking is thinking. Thus the deceiver hypothesis here works, so to speak, to set a boundary or limit to what can be doubted.¹⁷ It is as if Descartes were saying, "I can be

¹⁷ I owe the idea of "boundary-setting" to Jonathan Bennett.

fooled up to this point but not beyond it." To be sure, this "boundary-setting" aspect of the deceiver hypothesis is not a proof that "I am thinking" is certain, because one could not sensibly argue, "it is certain that if I'm deceived into thinking that I'm thinking, then I am thinking; it is certain that I am deceived into thinking that I'm thinking; therefore it is certain that I am thinking." Rather, the "boundary-setting" aspect of the deceiver hypothesis works negatively, by showing that even this most radical skeptical hypothesis cannot *disprove* the certainty of "I am thinking."

Having applied the method of doubt to his former beliefs about himself, Descartes introduces, in step 8, his revolutionary conception of the self as essentially a spiritual (mental) substance, mind, or soul. It is easy to miss the revolutionary nature of Descartes's conception; for, whether or not you personally believe in the existence of a soul that could continue to exist without the body, you can surely recognize this notion of the soul as central to the Judeo-Christian view of human beings. The conception of a human being as one composed not only of a physical part (the body) subject to all the laws of biology and physics but also a nonphysical part (the soul) not limited by these laws, whether or not you personally *accept* it, is at least very *familiar* to you; for it is the modern version of the Judeo-Christian view of humans. But when Descartes wrote his *Meditations*, the prevailing conception of the soul was the Aristotelian one just sketched, according to which the soul and the body form a single substance. This is not to say that Descartes's conception of the self was totally unprecedented in the history of philosophy. On the contrary, Plato had conceived the soul as an immaterial entity that outlives the body, as is clear, for instance, from his Dialogue *Phaedo*. And there are important vestiges of Plato's view in Aristotle—and even more so in Aquinas and other medieval Christian thinkers. But what Descartes did was renovate Plato's dualistic conception of the self by giving an extraordinarily sharp, clear account of it—one that he intended to be satisfactory even in light of the scientific revolution of the seventeenth century—as well as forceful new arguments for it. His account, then, provides the modern philosophical underpinnings for the Judeo-Christian view of human beings.

Descartes was aware that his conception of the self was bound to seem barren, uninformative, and overly abstract to his readers. So in *Meditation II*, he does two different things to combat this impression. Let us look at them in turn.

First, Descartes tries to give content to his conception of a purely thinking substance by reminding us of what “thinking” covers—namely, all conscious states. These include doubting, understanding, asserting, denying, willing, refusing, imagining, and seeming to perceive. As Descartes puts it, ingeniously listing the very thoughts he has been reporting in the *Meditations*:

This is a considerable list, if everything on it belongs to me. But does it? Is it not one and the same ‘I’ who is now doubting almost everything, who nonetheless understands some things, who affirms that this one thing is true, denies everything else, desires to know more, is unwilling to be deceived, imagines many things even involuntarily, and is aware of many things which apparently come from the senses? (CSM II 19, SPW 83, AT VII 28)

Second, Descartes tries to show that the conception of a physical object, though it initially seems easier to grasp than the conception of a pure mind, is in fact just as abstract. As one writer insightful notes, this is the purpose (or, rather, one of the purposes) of the passage about the wax.¹⁸ This can be seen by attending to the way Descartes introduces that passage:

From all this I am beginning to have a rather better understanding of what I am. But it still appears—and I cannot stop thinking this—that the corporeal things of which images are formed in my thought [by imagination], and which the senses investigate, are known with much more distinctness than this puzzling ‘I’ which cannot be pictured in the imagination. And yet it is surely surprising that I should have a more distinct grasp of things which I realize are doubtful, unknown and foreign to me, than I have of that which is true and known—my own self. But I see what it is: my mind enjoys wandering off and will not yet submit to being restrained within the bounds of truth. Very well then; just this once let us give it a completely free rein, so that after a while, when it is time to tighten the reins, it may more easily submit to being curbed.

¹⁸ Arthur Danto, *What Philosophy Is*, p. 93.

Let us consider the things which people commonly think they understand most distinctly of all; that is, the bodies which we touch and see. I do not mean bodies in general—for general perceptions are apt to be somewhat more confused—but one particular body. Let us take, for example, this piece of wax. (CSM II 20, SPW 83–4, AT VII 29–30)

Having carefully prepared us for a comparison of his novel conception of the self (which can be neither perceived by the senses nor pictured by imagination) with the seemingly easier conception of an ordinary physical object (which can be examined by the senses and pictured by imagination), Descartes asks, “So what was it in the wax that I understood with such distinctness?” (CSM II 20, SPW 84, AT VII 30) This question concerns our *conception* of the wax; it can be paraphrased as, “What constitutes our conception of a body, for example, this piece of wax?” Descartes arrives at his answer by a process of elimination. First eliminated are the wax’s observable properties—its shape, size, texture, color, smell, and so on. These do not constitute our conception of the wax, because even if they all change, the wax remains. So our conception must be an abstract one: “Let us concentrate, take away everything which does not belong to the wax, and see what is left: merely something extended, flexible, and changeable” (CSM II 20, SPW 84, AT VII 31). What Descartes says next makes the conception even more abstract. He asks, Since we conceive the wax as something “flexible” and “changeable,” is not our conception of the wax composed of the specific sizes and shapes that we can *imagine* the wax taking on? No, he answers; for we conceive that the wax can take on infinitely many different shapes and sizes, but we can only imagine (i.e., picture or visualize) a finite number of these. At last, Descartes turns his attention to “extended”:

And what is meant by ‘extended’? Is the extension of the wax also unknown? For it increases if the wax melts, increases again if it boils, and is greater still if the heat is increased. I would not be making a correct judgement about the nature of wax unless I believed it capable of being extended in many more different ways than I will ever encompass in my imagination. (CSM II 21, SPW 84, AT VII 31)

In this important sentence, we may interpret Descartes as saying that to conceive the wax as something extended is to conceive it as something

that *can take on* a great many different shapes and sizes (something three-dimensional). Our conception of the wax, then, is merely the conception of something that can take on various (three-dimensional) shapes and sizes. Descartes adds that we have this conception neither by the senses nor by imagination but only by reason. Moreover, it is now a “clear and distinct” conception, unlike the imperfect and confused one he began with; for, unlike a conception involving the properties perceived by the senses or involving the specific shapes and sizes pictured by imagination, his conception now contains all that—and only what—the wax must have to remain the same wax. It is a purified conception, arrived at by a careful process of reasoning.

Finally, notice how Descartes has fulfilled his purpose of showing that our conception of a body is as abstract as his conception of a mind. The purified, “clear and distinct” conception of a body, as we’ve just seen, is the conception of something that *can take on various shapes and sizes*. But what is the conception of a purely thinking substance but the conception of something that *can take on various thoughts (cognitiones)*—doubts, desires, beliefs, sensations, and so on? Moreover, neither conception comes from the senses or imagination; both are purely intellectual. The upshot is that Descartes’s radically novel conception of the self as a purely thinking substance is no more abstract or difficult than the true conception of a material body.

10. Cartesian Dualism

By the time Descartes has finished discussing the piece of wax, he has, in effect, expounded his single most characteristic doctrine, namely, Cartesian Dualism. Cartesian Dualism is the view that the universe contains two radically different kinds of substance: (1) mind, defined as a *thinking, unextended* substance, (2) body (i.e., matter), defined as an *extended, unthinking* substance. Notice, then, that mind and matter are defined as opposites: mind is thinking, while matter is unthinking; matter is extended (three-dimensional) while mind is unextended. Matter occupies space but doesn’t think, and a mind thinks but doesn’t occupy space. When trying to focus on Descartes’s distinctive concepts of mind and matter, it is helpful to bear in mind his Latin names for them: *res cogitans* (“thinking thing”) and *res extensa* (“extended thing”).

Descartes’s dualism is perhaps his most important doctrine, for it had sweeping and far-ranging implications. On the one hand, it helped to

clear the way for modern physical science. As we saw in the last chapter, the prevailing, Aristotelian physics of Descartes's day held that the universe is inherently purposeful or teleological. In other words, everything that happened, whether it was the motion of the stars in the supralunar region or the growth of a tree in the sublunar region, was supposed to be explained by certain purposes, goals or ends working themselves out within nature. Aristotle called such purposes "final causes," and final causes were considered to be indispensable for explaining the ways nature operated. A battle was taking shape between the scholastic defenders of this traditional science and the proponents of the new science of Kepler and Galileo, which denied the relevance of final causes for explaining nature. Descartes's dualism provided a powerful philosophical rationale for the newer conception, for one implication of Descartes's dualism is that final causes are expelled from the physical universe, or *res extensa*. The only place left for final causes is the mind, or *res cogitans*. Thus, Descartes's Dualism helped prepare the way for modern physics, which does not explain nature by reference to purposes. On the other hand, Cartesian dualism also ensured the possibility of immortality; for if the mind or soul is really a different substance from the body, then the destruction of the body does not entail the extinction of the mind. Thus, Cartesian Dualism simultaneously helped to clear the way for modern physics and held the door open for religious beliefs about the immortality of the soul.

Before leaving *Meditation II*, we need to raise a question: How far has Descartes really come toward establishing his dualism at this point in his *Meditations*?—for although Descartes, by the end of *Meditation II*, has *explained* his dualism by expounding his conceptions of both *res cogitans* and *res extensa*, it doesn't follow that he has shown this dualism to be true, that is, shown that mind and matter really are two different substances, one thinking and unextended, the other extended and unthinking. How close has Descartes really come to proving this at the end of *Meditation II*?

To begin with an obvious point, he certainly has not fully established his dualism, because he has not shown that *there is* such a thing as matter. Remember that Descartes still doubts the existence of the material world. So he certainly hasn't shown (and isn't claiming to have shown) that the universe actually *contains* any *res extensa*. Not until the sixth (and last) *Meditation* does he try to show this.

The significant question, rather, is whether Descartes has already shown that mind is a *different* substance from any matter that *may* exist,

if any exists. In other words, has he shown that there is such a thing as a purely thinking substance? This is a more delicate issue, which Descartes's contemporaries repeatedly asked him to clarify.

To focus the question better, let's leave the *Meditations* for a moment and consider a passage from Descartes's *Discourse on the Method*, Part Four, paragraph 2:

Next I examined attentively what I was. I saw that while I could pretend that I had no body and that there was no world and no place for me to be in, I could not for all that pretend that I did not exist. I saw on the contrary that from the mere fact that I thought of doubting the truth of other things, it followed quite evidently and certainly that I existed; whereas if I had merely ceased thinking, even if everything else I had ever imagined had been true, I should have had no reason to believe that I existed. From this I knew I was a substance whose whole essence or nature is simply to think, and which does not require any place, or depend on any material thing, in order to exist. Accordingly this "I"—that is, the soul by which I am what I am—is entirely distinct from the body . . . and would not fail to be whatever it is, even if the body did not exist. (CSM I 127, SPW 36, AT VI 32–3)

Here, Descartes appears to be saying that, merely from the fact that he can doubt the existence of matter but cannot doubt his own existence, it follows that he is a purely thinking substance. There is a similar passage in his *Search After Truth*: "I . . . am not a body. Otherwise, if I had doubts about my body, I would also have doubts about myself, and I cannot have doubts about that" (CSM II 412, AT X 518). Descartes's line of reasoning in such passages has been called the "Argument from Doubt." It may be summarized this way:

- (1) I can doubt that (my) body exists.
- (2) I cannot doubt that I exist.

. . . I am not a body.

Is this argument valid? No, it isn't. Just because I can doubt that my body exists but not that I exist, it doesn't follow that I am not a body; for I might very well *be* a body but not *know* it. The Argument from Doubt is no better than this argument (which we can imagine Louis XVI giving

before the French revolution that ended his reign): “I can doubt that the last King of France exists; I cannot doubt that I exist; therefore I am not the last King of France.” Sometimes the fallacy in such arguments is called the “masked man fallacy”: “I know who my father is; I do not know who the masked man before me is; therefore this masked man is not my father.”

Does Descartes rely on the Argument from Doubt in his *Meditations*, as he appears to have done in his *Discourse on the Method* and *Search After Truth*? It would be unfortunate if he did, for the *Meditations* are Descartes’s most careful and rigorous presentation of his philosophy.

Initially, it may *look* as if Descartes uses the Argument from Doubt even in his *Meditations*. For in *Meditation II* he does say:

I am not that structure of limbs which is called the human body. I am not even some thin vapor which permeates the limbs—a wind, fire, air, breath, or whatever I depict in my imagination; for these are the things which I have supposed to be nothing. Without changing this supposition, I find that I am still certain that I am something. (CSM II 18 with n. 3, SPW 82 with n. 3, AT VII 27)

Matters of detail apart, this is again the Argument from Doubt. The transition from step 7 to step 8 in our summary of the *Meditation* also looks much like the Argument from Doubt.

The very next sentence after the passage just quoted, however, shows that Descartes does not wish to rely on that argument in his *Meditations*, for he says:

And yet may it not perhaps be the case that these very things which I am supposing to be nothing, because they are unknown to me, are in reality identical with the ‘I’ of which I am aware? I do not know, and for the moment I shall not argue the point, since I can make judgements only about things which are known to me.

Here Descartes pulls back from asserting that he is not a body; he admits that for all he knows at this point, he *may* be one. All he knows is that he is *at least* a thinking being.

Accordingly, step 8 in our summary of *Meditation II*—the claim that “I am only a thing that thinks”—should not be understood to mean “I know

that I am only a thing that thinks." Rather, step 8 means "I know only that I am a thing that thinks." The position of the word "only" is crucial. This word serves to limit what Descartes is claiming to *know* about what he is, not to exclude body from what he may, in reality, be. The upshot is that the weakness of the Argument from Doubt does not vitiate Descartes's reasoning in *Meditation II*, simply because Descartes does not here rely on that argument. Not until *Meditation VI* does Descartes give an argument intended to prove definitively that the mind is a different substance from the body. We must wait until we reach that argument to decide whether it succeeds in establishing Cartesian Dualism.

Meditation III

The Criterion of Truth and the Existence of God

1. Descartes's Criterion of Truth

At the outset of *Meditation III*, Descartes summarizes the knowledge that he has attained in the previous two *Meditations*.

I am a thing that thinks: that is, a thing that doubts, affirms, denies, understands a few things, is ignorant of many things, is willing, is unwilling, and also which imagines and has sensory perceptions; for, as I have noted before, even though the objects of my sensory experience and imagination may have no existence outside me, nonetheless the modes of thinking which I refer to as cases of sensory perception and imagination, insofar as they are merely modes of thinking, do exist within me—of that I am certain.

In this brief list I have gone through everything that I truly know, or at least everything that I have so far discovered that I know.
(CSM II 24, SPW 87, AT VII 34–35)

So far, this is *all* that Descartes claims to know; everything else is still subject to doubt. Only what Descartes takes himself to have established by the *cogito*, that he thinks (i.e., doubts, affirms, denies, etc.) and therefore exists, is secure. So we may summarize his first step in *Meditation III* simply like this:

1. So far I know only that I am a thing that thinks.

In his remaining *Meditations*, Descartes will build upon this one unshakeable piece of knowledge, which is itself based solely on the *cogito*.

We have seen, of course, that the *cogito* is by no means as simple or unproblematic as it seems. But from this point on, let us put aside any worries about the *cogito*. For even if, despite the defense of the unreconstructed *cogito* that was offered in 2.7 and 2.8, you reject the *cogito* as a proof of one's own existence (on the ground that it is either question-begging or invalid), the certainty of "I am thinking" and of other *cogitationes*, the certainty of the inference from "I am thinking" to "I exist," and the certainty of "I exist," all remain secure. And these three points are sufficient for the further uses to which Descartes will put the *cogito*.

Having affirmed his existence as a thinking thing, Descartes now asks himself, Can I learn anything more from this? As he puts it:

Now I will cast around more carefully to see whether there may be other things within me which I have not yet noticed. I am certain that I am a thinking thing. Do I not therefore also know what is required for my being certain about anything? (CSM II 24, SPW 87, AT VII 35)

Descartes's idea here is this. He has one instance of absolutely certain, unshakeable knowledge. So, by examining this single, shining example of genuine knowledge, he should be able to discover the *feature* of it that makes it so unshakeable. He can then consider whether any further propositions also have this feature: if they do, he will be able to extend his knowledge to them, as well. We may summarize his idea this way:

2. Can I learn anything more from this? I am certain that I am a thinking thing. Don't I therefore also know what is required for my being certain about anything?

Accordingly, Descartes asks himself: what assures me that I am a thinking thing? What is the characteristic of this piece of knowledge that makes it so certain? He answers,

In this first item of knowledge there is simply a clear and distinct perception of what I am asserting. (CSM II 24, SPW 87, AT VII 35)

The feature of his "first item of knowledge" that renders it so certain, Descartes declares, is simply that it is "a clear and distinct perception."

So, he reasons, perhaps he can now safely generalize that *whatever* he perceives “clearly and distinctly” is true; for, as he continues:

[T]his would not be enough to make me certain of the truth of the matter if it could ever turn out that something which I perceived with such clarity and distinctness was false. So I now seem to be able to lay it down as a general rule that whatever I perceive very clearly and distinctly is true. (CSM II 24, SPW 87, AT VII 35)

Here, Descartes has extracted from his “first item of knowledge” his famous criterion of truth—that whatever he perceives “clearly and distinctly” is true. Let us more closely examine both (i) what he extracts this criterion *from* and (ii) the *content or meaning* of the criterion itself.

Descartes speaks as if the “first item of knowledge” from which he extracts his criterion is just the one clearly and distinctly perceived proposition, “I am a thinking thing.” Descartes’s knowledge that he is a thinking thing, however, is really a complex piece of information. Its elements include at least the various components of the basic, unreconstructed *cogito*, namely, the knowledge (a) that he is thinking, (b) that his thinking necessitates his existence, and (c) that he exists. We may interpret him, therefore, as saying that what assures him that he is a thinking thing is that he very clearly and distinctly perceives the *cogito*—here seen as a tight “package” of certainties composed of (a), (b), and (c). In other words, we can understand him as deriving or extracting his criterion of truth from the *cogito*, by means of the following argument: “If something could be clearly and distinctly perceived yet false, then this would cast a shadow of doubt on the *cogito* itself. But the *cogito* is absolutely indubitable. Therefore, what is clearly and distinctly perceived cannot be false; so it must be true.” Accordingly, we may summarize Descartes’s third step—his extraction of the “clarity and distinctness” criterion of truth from the *cogito*—as follows:

3. What assures me that I am a thinking thing? Only that I perceive the *cogito* very clearly and distinctly. So, it seems that I can already infer that whatever I perceive clearly and distinctly is true; for

(1) If my clear and distinct perceptions could be false, then the *cogito* would not be certain.

(2) The *cogito* is certain.

∴(3) My clear and distinct perceptions cannot be false; that is, whatever I perceive clearly and distinctly is true.

Adopting a recent commentator's usage, we can call this argument from (1) and (2) to (3), "the extraction argument."¹ That Descartes advances the extraction argument is confirmed by the following passage from the *Discourse on the Method*, Part Four, paragraph 4, where he explicitly derives his criterion of truth from the *cogito*:

After this I considered in general what is required of a proposition in order for it to be true and certain; for since I had just found one that I knew to be such, I thought that I ought also to know what this certainty consists in. I observed that there was nothing at all in the proposition 'I am thinking, therefore I exist' to assure me that I am speaking the truth, except that I see very clearly that in order to think it is necessary to exist. So I decided that I could take it as a general rule that the things we conceive very clearly and distinctly are all true. (CSM I 127, SPW 36, AT VI 33)

But what exactly does Descartes mean by a "clear and distinct perception"? In his *Principles of Philosophy*, under the caption "What is meant by a clear perception, and by a distinct perception," he offers the following definition:

I call a perception "clear" when it is present and accessible to the attentive mind—just as we say that we see something clearly when it is present to the eye's gaze and stimulates it with a sufficient degree of strength and accessibility. I call a perception "distinct" if, as well as being clear, it is so sharply separated from all other perceptions that it contains within itself only what is clear. (CSM I 207–8, SPW 174–5, AT VIIIA 22)

This definition is not as helpful as one might like, for it is basically just an analogy: a clear and distinct perception is an intellectual perception *like*

¹ Gary Hatfield, *Descartes and the Meditations*, p. 180.

the visual perception of an object in good conditions, when we can discriminate the object from its environment and make out each of its various parts. This does not give us a definite way of telling when a perception is clear and distinct; and, indeed, Descartes admits in a number of places that this is not always easy to do.

One standard interpretation of Descartes's notion of clarity and distinctness sees it as being inspired by mathematics. The idea is that a simple mathematical proposition, like $1 + 1 = 2$, is so clear and obvious that it cannot be doubted and that it is propositions of this kind that Descartes calls "clear and distinct." As Frederick Copleston puts it in his well-known, multi-volume *History of Philosophy*:

This criterion of truth was doubtless suggested to Descartes by mathematics. A true mathematical proposition imposes itself, as it were, on the mind: when it is seen clearly and distinctly, the mind cannot help assenting to it. Similarly, I affirm the proposition, *I think, therefore I am*, not because I apply some extrinsic criterion of truth, but simply because I see clearly and distinctly that so it is.²

Although this interpretation of Descartes's criterion is no doubt partly right, it is not completely satisfactory; for Descartes, as we have seen, derives the criterion from the *cogito*. But even if we regard the *cogito* in its simplest, classic, unreconstructed form (as we are now doing), it has a certain complexity—a complexity that gets masked when the *cogito* is referred to as "one" proposition (as it is by Descartes himself in the passage we quoted above from his *Discourse on the Method*) or as "the proposition, *I think, therefore I am*" (as it is by Copleston in the passage just cited); for the unreconstructed *cogito* has three components: "I am thinking," "If I am thinking, then I exist" (or "it is impossible for what thinks not to exist"), and "I exist." (The generalization "it is impossible for what thinks not to exist," as we suggested in 2.8, is grasped only by first grasping the necessary connection between thinking and existing in one's own case. This premise, though it is needed for the argument's formal validity, is recognized in retrospect; it is not needed to grasp the blinding force of the *cogito*. Descartes would presumably say the same

² Frederick Copleston, *A History of Philosophy*, p. 98.

thing about the conditional “if I am thinking, then I exist”: it is known only as a result of grasping in one’s own case that one must exist since one is thinking.) Furthermore, the certainty of each of these components stems from a different source. “I am thinking” is certain because it reports only my own present conscious state. “If I am thinking, then I exist” is certain because it is obviously impossible for “I am thinking” to be true and “I exist” to be false. “I exist” is certain because it is the conclusion of an argument which has the features that (a) one of its premises reports only my own present conscious state and self-evidently necessitates the argument’s conclusion, and (b) it becomes formally valid with the addition of a premise that expresses the necessity referred to in (a).

Now, in order for clear and distinct perceptions all to be akin to mathematical ones, Descartes’s criterion would have to be derived exclusively from the second component—from the proposition “if I am thinking, then I exist” (or “it is impossible for what thinks not to exist”); for only this proposition shares with mathematical propositions the feature that primarily explains their certainty, namely, that a true mathematical proposition is *necessarily* true. The proposition that $1 + 1 = 2$, for example, doesn’t just happen to be true. Rather, it must be true; it could not possibly be false. As philosophers usually put it today, there are no *possible worlds* in which one plus one does not equal two. The same goes for the proposition “if I am thinking, then I exist”: this proposition couldn’t possibly be false. There are no possible worlds in which I am thinking but I do not exist. By contrast, the propositions “I am thinking” and “I exist” could have been false, for I might not have existed. *There are* possible worlds in which I do not exist; and if any of those worlds had been actual, then “I exist,” as well as “I am thinking,” would have been false. Of course, if “I am thinking” or “I exist” are uttered or entertained in a given possible world, call it W, then those propositions must be true in W. But this does not mean that “I am thinking” or “I exist” are true in all possible worlds, for they need not be uttered or entertained in all possible worlds. So these propositions are not necessary; instead, they are *contingent*. (The term “contingent,” as it is used in philosophy, applies to all propositions that are neither necessary nor impossible.) Yet, Descartes evidently regards “I am thinking” and “I exist” as “clear and distinct,” too. So his criterion can be satisfied by propositions of fundamentally different types—contingent ones as well as necessary ones. Furthermore, Descartes surely regards the validity of the argument

If I am thinking, then I exist
I am thinking

. . . I exist

as “clear and distinct.” So his criterion can also be satisfied by the logical step from the premise(s) of a valid argument to its conclusion. Finally, Descartes regards “I exist” as clear and distinct; so his criterion can be satisfied by the conclusion of a valid argument whose premises describe only one’s own present thoughts or are obvious necessary truths.

The upshot is that, in addition to being satisfied in (1) the special case of a proposition (“I exist”) that is self-evidently necessitated by a proposition that describes only one’s own present thoughts (“I am thinking,” or any other *cogitatio*-statement), Descartes’s criterion of clarity and distinctness can be satisfied by items of four different kinds: (2) contingent propositions describing only one’s own present thoughts, (3) obvious necessary propositions, (4) the step, transition, or inference from the premise(s) to the conclusion of a valid argument, and (5) the logical consequences of (2)’s and/or (3)’s. No doubt, this is what Descartes intended; for, as we shall see, he goes on to use his criterion throughout his attempt to rebuild his knowledge: the criterion has, so to speak, a lot of work to do. Perhaps the best way to interpret Descartes’s criterion, then, is to see it as a kind of “pass,” “ticket,” or “license,” saying that in rebuilding one’s knowledge, it is permissible to build on the five types of items just listed.

Although Descartes’s criterion of truth may now seem quite complex and perhaps even a bit slippery, this need not worry us further; for nothing will stop us from asking whether the specific propositions that Descartes will put forward as “clear and distinct” are really as unquestionable as he takes them to be—whether they are as obviously certain as the *cogito* and its component parts.

Let us conclude this section, then, by summarizing how Descartes will use his criterion of truth in his subsequent *Meditations*. Basically, he will use it to overcome his doubt concerning all matters beyond his own existence as a thinking thing. Specifically, he will use it to show:

1. that a perfect God exists (*Meditation V*),
2. that mind is really a different substance from any matter that may exist (*Meditation VI*), and
3. that the material world exists (*Meditation VI*).

Descartes's basic strategy, then, is to derive his criterion of truth from the *cogito* and then to apply it in successively proving God's existence, mind-body dualism, and the existence of the physical world.

2. The Project of *Meditation III*

The three-point strategy just outlined highlights the three main theses that Descartes will seek to establish in his subsequent *Meditations*: the existence of God, mind-body dualism, and the existence of the physical world. However, as you may surmise from the fact that it skips over the rest of *Meditation III* and *Meditation IV*, it is an oversimplified account of Descartes's strategy; for even after extracting his criterion of truth from the *cogito*, Descartes is not yet willing to use it in arguing for his remaining three main theses. Instead, he embarks on a fascinating, famous, and (as we shall see) problematic attempt to justify or vindicate that very criterion. The key element in this attempt, as we shall see, is a complex argument for the existence of a perfect God who guarantees the truth of clear and distinct perceptions (thus the argument for God's existence in *Meditation V* is not the first such argument that Descartes gives). It is to Descartes's quest for a vindication of his criterion of truth, then, that we must next turn our attention. We shall initially proceed, as before, by summarizing his reasoning step-by-step.

In the paragraph immediately following the extraction argument, Descartes reminds himself that much remains doubtful. We may paraphrase him this way:

4. I must remember my doubt concerning many things whose existence once seemed obvious, that is, physical objects that I perceived by my senses. Only the fact that I have ideas of these objects is clearly and distinctly perceived. The origin of these ideas and whether they resemble anything existing outside me are still unknown.

The next paragraph is an extremely crucial one. In its first segment ("But . . . mind's eye"), Descartes turns his attention to the most obvious propositions he can identify, such as simple mathematical ones. We can paraphrase the segment this way:

- 5. But can't I now at least say that I perceive simple propositions of mathematics, such as $2 + 3 = 5$, clearly enough to affirm their truth? The only reason I've found for doubting such things is that perhaps God gave me a nature such that I am deceived even about what seems most evident. Whenever I think of an all-powerful God, I must admit that if he wishes, he can easily make me go wrong even about things that I think I perceive utterly clearly.**

Here, then, Descartes is expressing a doubt even about propositions that he most clearly and distinctly perceives—a doubt based on the possibility of a deceiving God that he raised in *Meditation I*. But now, consider what he says in the next segment (“Yet . . . contradiction”). We may paraphrase it this way:

- 6. But every time I actually attend to these things, I am so convinced by them that I'm impelled to say: “No one can cause me not to exist so long as I am thinking that I exist, or make it true that I've never existed since I now exist, or even that two plus three are not equal to five; or that any other proposition in which I see an obvious contradiction is true.”³**

Here, Descartes asserts that *at the actual time* that he is clearly and distinctly perceiving the *cogito*, or a simple instance of the law of noncontradiction (i.e., “I cannot both never exist and exist now”), or even a simple mathematical proposition, he *cannot doubt it*. Such propositions as *I am thinking, therefore I exist, or not both p and not-p*, or even $2 + 3 = 5$ are, to borrow a term coined by E. M. Curley, “assent-compelling.”⁴ They cannot be doubted *during the time that one is clearly and distinctly perceiving them*.

But doesn't this contradict what Descartes said in the preceding segment, where he admitted that an omnipotent God could deceive him even about things he most clearly and distinctly perceives? No, for there Descartes did not admit that he could doubt a proposition *while clearly and distinctly perceiving it*. Rather, he admitted that *while thinking about an*

³ The need for the word “even,” which does not appear in John Cottingham’s translation of the Latin text of the segment that is here being closely paraphrased, was pointed out to me in correspondence by Jean-Marie Beyssade, as the correct translation of Descartes’s “*vel forte etiam*,” which Cottingham translates, like “*vel*,” simply as “or.”

⁴ Edwin M. Curley, *Descartes Against the Skeptics*, p. 119.

omnipotent God, he had to concede that such a God would be able to deceive him even about the most obvious things. From this concession, it does not follow that Descartes can doubt any one of those things *while actually paying attention to it*. It only follows that even though he cannot doubt a proposition so long as he clearly and distinctly perceives it, he can doubt whether the fact that he clearly and distinctly perceives a proposition guarantees that it is true. As James Van Cleve succinctly puts it: “[Descartes] might be uncertain of the general connection between clear and distinct perception and truth, yet certain of every proposition [he] . . . clearly and distinctly perceive[s].”⁵ Thus, in the first two segments of the paragraph (points 5 and 6 of our paraphrases), Descartes is weighing his certainty about particular, occurrent clear and distinct perceptions, against a general doubt concerning the reliability of his cognitive faculties (notably his faculty for clear and distinct perception), based on the possibility of a deceiving God.

In the third and final segment of the paragraph (“And since . . . else”), Descartes declares that there is only one way he can emerge from this oscillation between doubt and certainty. We can paraphrase the segment this way:

7. Although my reason for doubting clear and distinct perception is very slight and “metaphysical,” to remove it I must determine whether (a) God exists, and (b) God can be a deceiver; for without knowing this, I cannot be perfectly certain of anything else.

Thus, in order to resolve the tension between his certainty about particular, occurrent clear and distinct perceptions and his doubt about the reliability of clear and distinct perception in general, Descartes believes he must eliminate the possibility of a deceiving God, by establishing the existence of a nondeceiving God. He must show that his clear and distinct perceptions are guaranteed to be true by God himself.

3. From the Idea of God to God

Accordingly, Descartes now turns to the question of God’s existence, which becomes the main topic of *Meditation III*. In that *Meditation*, Descartes advances two related proofs of God’s existence. (He gives a third,

⁵ James Van Cleve, “Foundationalism, Epistemic Principles, and the Cartesian Circle,” p. 67.

quite different proof—the famous Ontological Argument—in *Meditation V*). Both *Meditation III* proofs are Descartes's own special versions of what is called the Cosmological Argument for the Existence of God. The Cosmological Argument, which comes in several different versions, attempts to prove that God exists by showing that the existence of the world, or of things in the world, requires an original cause or an ultimate explanation. (By contrast, the Ontological Argument given in *Meditation V* attempts to prove that God exists because of the very concept or definition of God as an absolutely unsurpassable being.) For example, Saint Thomas Aquinas's seminal versions of the Cosmological Argument in the thirteenth century reason from the existence of certain effects in the world to God as the first cause of those effects. Like Aquinas's arguments, Descartes's *Meditation III* proofs reason from certain effects to God as the cause of those effects. However, unlike Aquinas's arguments, Descartes's proofs cannot appeal to any of God's effects in the physical world; for remember that at this point in the *Meditations*, the existence of the entire physical world is still in doubt. Accordingly, Descartes's strategy is to argue from the *idea* of God that he finds in his mind to God as the cause of that idea. His first proof starts just from the idea of God and attempts to show that God himself is the only possible cause of that idea. His second proof builds on the first by trying to show that only God could cause the existence of a thinking thing that has the idea of God. The second proof depends upon the first, so we shall concentrate our attention on the first.

3.1 The Nature of Ideas

Since Descartes's argument starts from the idea of God, he prepares the way for the argument by discussing the nature of ideas. We shall simply list the points he makes, and comment briefly on each of them.

1. An idea represents something; it is like a picture.

This is Descartes's most fundamental and influential point about ideas. An idea, according to him, is essentially a mental representation of its object—of the thing *of* which it is an idea. This is true whether or not that object really exists, since an idea must have a content, must be an idea of *something*. The same view of ideas can be found in many major philosophers that Descartes influenced, including Locke, Hume, and Kant.

2. An idea itself cannot be false.

Descartes's point can be put this way: even if a picture doesn't depict accurately or if what it depicts doesn't exist, the picture itself—considered merely as an image—cannot be false (or true). Rather, the picture is just something that exists in its own right, whether or not what it depicts also exists. Likewise, an idea, considered purely in terms of its content, that is, without regard to whether that content corresponds to anything else, cannot be either false or true.

3. Falsity (and truth) becomes possible only when I make a judgment, especially when I judge that an idea in my mind corresponds to or resembles something outside my mind.

Given that an idea itself cannot be false (or true), truth and falsity become possible only when some *judgment* or *assertion* is made with respect to ideas. Descartes's basic point here, which is still a commonplace in philosophy, is that truth and falsity pertain not to concepts or ideas but, rather, to assertions, statements, judgments, propositions, and the like. For example, the concept or idea *horse* is neither true nor false. Only an assertion or proposition that uses this concept (e.g., "Some horses are thoroughbreds") can be true or false.⁶ Since Descartes's chief purpose after establishing his own existence is to attain knowledge of things existing outside his own mind (first God, then other things), he here

⁶ In a brief but difficult passage in the *Third Meditation*, Descartes complicates matters by saying that some ideas, specifically those of "light and colours, sounds, tastes, heat and cold, and the other tactile qualities" (i.e., of secondary qualities, see chapter 6, section 4.1) are "materially false." A rough but useful characterization of material falsity is this: "In the Third Meditation, some sensory ideas were said to be 'materially false,' which means that they provide material for false judgment (AT VII 43 [see also CSM II 162–64, AT VII 231–35]). (Such judgments occur when one affirms that . . . external objects have properties in them that resemble our sensations of color, sounds, and other so-called secondary qualities)" (Gary Hatfield, *Descartes and the Meditations*, p. 190). In the paragraph where Descartes introduces his view that ideas of color, heat and cold, sound and so forth are materially false, he contrasts these with ideas that, as when he examined the idea of the wax, he clearly and distinctly perceived to represent properties of material things (if such things exist, which is unknown at this point), including "size, or extension in length, breadth, and depth; shape, which is a function of the boundaries of this extension; position, which is a relation between various items possessing shape;

emphasizes the type of error he is most anxious to avoid—judging that some idea in his mind corresponds to a reality outside his mind when it does not.

Having touched on the basic issue still before him (i.e., the correspondence or noncorrespondence of the ideas in his mind to things existing outside his mind), Descartes now focuses this issue more sharply, by presenting a possible classification of ideas.

4. My ideas seem to fall into three classes:

- (a) innate (i.e., inborn),**
- (b) adventitious (caused by objects located outside me),**
- (c) fictitious (invented by me).**

It is important to understand that Descartes is not at this point affirming that his ideas actually fall into these three classes, but only that they “appear to” do so; for, remember, he does not yet know—but is only beginning to investigate—the causes of his ideas. Thus, he immediately adds, in effect,

5. I cannot yet be sure how my ideas really divide up; perhaps they all fall into just one of these three classes.

As he puts it: “But perhaps all my ideas may be thought of as adventitious, or they may all be innate, or all made up; for as yet I have not clearly perceived their true origin” (CSM II 26, SPW 89, AT VII 37–38).

Next, Descartes examines the reasons he formerly had for thinking that some of his ideas fall into class (b) and, further, that those ideas resembled the objects that caused them. In effect, he here puts up for examination a standard seventeenth-century view of perception that has adherents to this day, called the “representational theory of perception.”

and motion, or change in position” (CSM II 28, SPW 92, AT VII 43). This suggests that part of what Descartes also means by an idea’s being materially false is that its content exhibits features that are different from anything that belongs to material objects as they would be described (assuming they existed at all) in a correct (Cartesian) physics, and in that sense falsely portrays, or misleads us about, the nature of the world. Descartes scholars have proposed different interpretations of his difficult notion of material falsity, which will not be discussed in further detail in this book, since it does not play any role in his central arguments. More will be said in chapter 6, however, about Descartes’s physics.

According to this view, when a person perceives a material object, the object causes an idea in the person's mind, the person has absolutely certain knowledge of the idea, and the idea gives the person knowledge of the object by representing it, which it does by resembling at least some of its properties (see Figure 3.1).

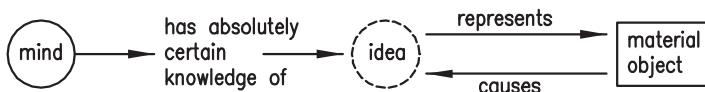


Figure 3.1

As we shall see later, Descartes himself eventually endorses such a view, but what he examines in *Meditation III* is his previous reasons for accepting it. We can paraphrase him this way:

6. My reasons for thinking that some of my ideas come from, and resemble, things existing outside me are (1) that “nature has taught me to think this” and (2) that the ideas occur independently of my will.

Descartes now finds these reasons to be very weak; for the first one only means that he has a natural, spontaneous inclination to believe that some ideas proceed from and resemble external objects. Descartes contrasts this inclination with the “natural light.” This is none other than the capacity for clear and distinct perception. He makes the strong claim that “whatever is revealed to me by the natural light . . . cannot in any way be open to doubt” (CSM II 27, SPW 89, AT VII 38)—a claim that should be seen in the context of the oscillation between doubt and certainty discussed earlier. But whatever may be said in favor of the natural light, the same cannot be said for “nature”; for just as natural impulses can drive one to choose evil over good, so they can lead one to choose error rather than truth. As for the fact that some ideas come independently of one’s will, it proves nothing; for perhaps some unknown faculty within the self produces those ideas anyway, much as happens in dreams. And even if the ideas did come from external objects, it would not follow that they must resemble those objects. Thus, Descartes concludes that,

7. These reasons are very weak; I must find “another way of investigating whether some of the things of which I possess ideas exist outside me” (CSM II 27, SPW 90, AT VII 40).

3.2 Objective Reality and Formal Reality

To grasp Descartes's way of investigating this question, we need to understand a metaphysical framework that he presents in the rest of the short but important paragraph from whose first sentence we have just quoted. The paragraph continues as follows:

In so far as the ideas are <considered> simply <as> modes of thought, there is no recognizable inequality among them: they all appear to come from within me in the same fashion. But in so far as different ideas <are considered as images which> represent different things, it is clear that they differ widely. Undoubtedly, the ideas which represent substances to me amount to more and, so to speak, contain within themselves more objective reality than the ideas which merely represent modes or accidents. Again, the idea that gives me my understanding of a supreme God, eternal, infinite, <immutable>, omniscient, omnipotent and the creator of all things that exist apart from him, certainly has in it more objective reality than the ideas that represent finite substances. (CSM II 27–28, SPW 90, AT VII 40)

Here Descartes begins by adding a further point to those already made about the nature of ideas. This new point is that ideas can be regarded in two different ways: (a) as states of the thinker or “modes of thought” and (b) as representations of their objects. An analogy may be helpful. Consider a fresco painted on a wall. The fresco can be regarded in two quite different ways. It can be regarded simply as an array of colors and shapes on the wall. Or it can be regarded as a representation of, say, Julius Caesar. Likewise an idea, since it is a mental representation of something, has two aspects. On the one hand, it is simply a state of the thinker—an episode or occurrence in the thinker’s mental history. On the other hand, it is a representation of its object. This distinction is nicely explained in the following passage by A. S. Pringle-Pattison:

It is important to remember . . . the distinction signalized by Descartes between an idea as a mental state, a psychical occurrence, and the same idea functioning in knowledge and conveying a certain meaning. . . . [I]n [the former] respect all ideas stand upon the same footing. . . . The treatment of ideas so regarded belongs to psychology. But ideas not only exist as facts

in the mental history of this or that individual; they have also . . . a “content” or meaning; they signify something other than themselves. We regard them, in Descartes’s words, “as images, of which one represents one thing and another a different thing,” and this is [an] important aspect of ideas.⁷

Having distinguished these two ways of regarding ideas, or two aspects of ideas, Descartes makes a further observation. Considered merely as states of a thinker, all ideas have the same status: they are just modes or (accidental) properties (of the thinker). But considered as representations of their objects, they do not all have the same status; for some of them represent other modes or properties; some represent finite substances, and one represents an absolutely infinite substance, namely, God. For example, my idea of squareness represents squareness, which is a mode or property of physical things, while my idea of fear represents fear, which is a mode or property of thinking things. Again, my idea of myself represents me, a finite thinking substance; while my idea of a stone represents a stone, which is a finite extended thing.⁸ Finally, my idea of God represents an infinite substance, namely, God. Diagrammatically, we can represent the metaphysical framework that Descartes has just introduced as shown in Figure 3.2.

Of course, at the beginning of *Meditation III*, Descartes is not asserting or assuming that all the objects of his ideas actually exist. On the contrary, at this point he is claiming to know only that one finite thinking substance and some modes of that substance exist—namely, himself and his own thoughts. But he is asserting that at least his *idea* of God, and his *ideas* of both thinking and extended finite substances and of their modes all exist. He is asserting that all the items on the left-hand side of our diagram exist and that a very few of the items on its right-hand side exist—namely, himself as a finite thinking substance and his own thoughts as modes or properties of that substance. Whether any of the other items on the right-hand side exist is still unknown, and is, indeed, the very issue Descartes is beginning to investigate.

⁷ John Locke, *An Essay Concerning Human Understanding*, p. 15.

⁸ We here abstract from the point, made in chapter 2, section 6.1, and to be revisited in chapter 6, that on Descartes’s final view a stone would not be a substance, but only a mode of an all-encompassing extended substance (which Descartes would still presumably regard as “finite” even if it were infinitely extended, because it would lack some of the attributes that belong only to the one truly infinite substance, namely God).

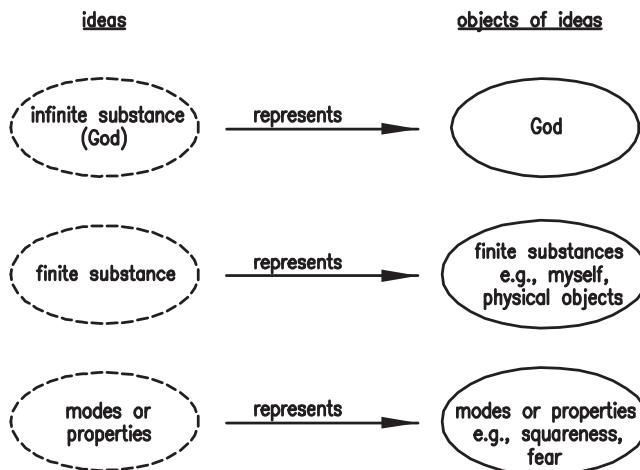


Figure 3.2

There is something else, however, that Descartes is asserting even at the beginning of *Meditation III*. He is asserting that the objects represented by his ideas—the things on the right-hand side of our diagram—have *different degrees of reality*: finite substances have more reality than modes, and an infinite substance has more reality than finite substances. You will ask how he can say this if he does not yet even know that any of those things (except himself and his thoughts) exist. The answer is that he is talking about the degree of reality that the things *would have if they existed*. Thus, Descartes's notion of "reality" must not be confused with actual existence. An analogy may be useful: even if all the money in the world were destroyed, it would still be true that \$1,000 is more than \$10. Likewise, Descartes is saying that even if no infinite substance exists, infinite substance has more reality than finite substance, and that even if no finite substance exists, finite substance has more reality than a mode. Notice also that Descartes frequently uses the term "more perfection" in place of "more reality": for him these terms are interchangeable. Historically speaking, this interchangeability, as well as the notion of degrees of reality, derive largely from Plato and the Neo-Platonists, for whom being (reality) and goodness are the same, and the highest good is also that which is most real.

Now on the one hand, the concept of degrees of reality may strike you as roughly right even though a bit fuzzy: you may be inclined to agree that in some sense, God has "more reality" than a mere finite substance and that a finite substance has "more reality" than a transient mode or

property. On the other hand, the concept of degrees of reality may seem obscure and questionable: you may want to protest that it makes no sense to talk about degrees of reality—that reality is an all-or-nothing matter, not one that admits of “more” or “less.” More will be said later about the basis for Descartes’s concept of degrees of reality. But for now, let us accept this concept, at least provisionally, for the sake of understanding how Descartes’s argument for the existence of God is supposed to work; for the concept of degrees of reality lies behind two key ideas in the argument, without which the argument cannot even be formulated.

The first key idea is that *some ideas represent their objects as having more reality than other ideas represent their objects as having*.⁹ Thus, the idea of a finite substance represents its object as having more reality than the idea of a mode represents its object as having, and the idea of God represents its object—God—as having more reality than the idea of a finite substance represents its object as having. The second key idea, which actually provides the basis for the first, is that *the degree of reality that an idea represents its object as having depends on the degree of reality possessed by the object itself*. In other words, if X has more reality than Y does, then the idea of X represents X as having more reality than the idea of Y represents Y as having. So, for example, since God has more reality than any finite substance, the idea of God represents (portrays) him as having more reality than the idea of a finite substance represents it as having.

To express these two key ideas, Descartes uses a technical terminology that he borrowed from medieval scholasticism and adapted for his own purpose. He puts the first key idea this way: *Some ideas contain more “objective reality” than others*. This claim means exactly the same thing as *Some ideas represent their objects as having more reality than other ideas represent their objects as having*. Thus, the term “objective” has here a completely different meaning from the modern one, where it has to do with objectivity—with what is actually the case independently of our beliefs and prejudices. You must erase all such connotations from your mind in order to understand Descartes’s meaning for this term. Instead, try to link the term “objective reality” with the notion of an *object of thought*; for an idea’s objective reality depends strictly on (the degree of reality possessed by) the object of the idea, on what the idea is *about*. It pertains to the second of the two ways of regarding ideas mentioned above—to their nature as representations of

⁹ The notion of an idea’s representing its object as having a certain amount of reality comes from Curley, *Descartes Against the Skeptics*, p. 126.

their objects. This has nothing to do with objectivity in the modern sense. It is also helpful to paraphrase Descartes's term in various ways. For example, in addition to "Some ideas represent their objects as having more reality than others," one could say, "Some ideas exhibit more reality in their contents than others." Some translators of the *Meditations* have chosen to substitute a completely different term, less misleading to the modern ear than "objective reality." For example, the British philosophers Elizabeth Anscombe and Peter Geach translate "objective reality," very aptly, as "representative reality."¹⁰ Using this translation, Descartes's point would be put this way: some ideas contain more representative reality than others.

Descartes's second key idea, we saw, is that the degree of reality an idea represents its object as having depends on the degree of reality possessed by the object itself (i.e., the degree of reality the object would have if it existed). To use the new term just introduced, this would be put by saying that an idea's degree of objective reality depends on the degree of reality had by the idea's object. To express this second key idea, however, Descartes introduces a second technical term, "formal reality." His idea, expressed using this term, in addition to "objective reality," is that *an idea's degree of objective reality depends on its object's degree of formal reality*. The term "formal reality" is, of course, just as new to you as was the term "objective reality." But at least the term "formal" does not have the misleading connotations that "objective" has. Think of formal reality as the kind of reality that a thing has, not in virtue of what it represents (that would be objective reality again), but, rather, in terms of its status as either a mode or property, a finite substance, or an infinite substance. Formal reality, then, is not too far removed from what people usually mean simply by "reality." It refers to a thing's actual status in the world—or at least to the status it would have if it existed. (Anscombe and Geach translate it as "actual or inherent reality.") Notice, then, that while objective reality is a special type of reality that belongs *only* to ideas in virtue of their representational function, everything, including ideas, has some degree of formal reality. Indeed ideas, being modes or properties of a thinker, have (along with modes or properties of other substances) the lowest degree of formal reality in Descartes's three-level hierarchy (modes, finite substances, and infinite substance).

We can now represent this hierarchy diagrammatically as shown in Figure 3.3.

¹⁰ René Descartes, *Descartes: Philosophical Writings*, pp. 81–82, 116.

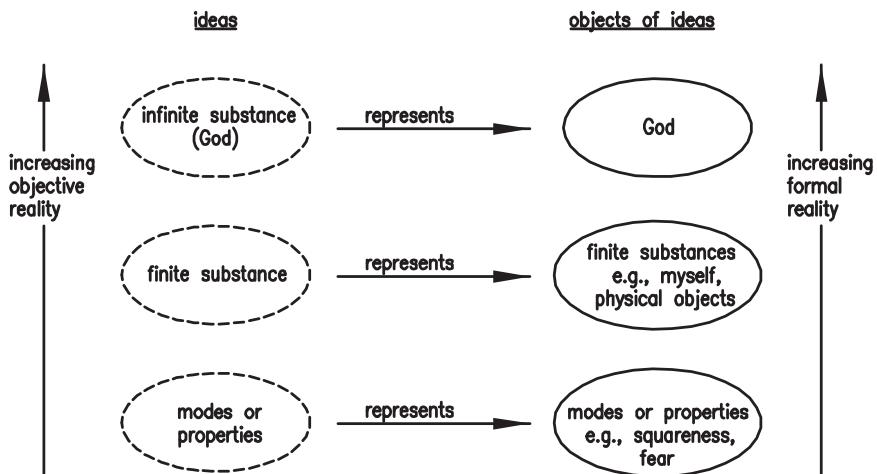


Figure 3.3

Note that the diagram positions ideas only in terms of their objective reality. In terms of formal reality, their position is at the right-hand bottom of the diagram: like fear or squareness, an idea is a mode or property of a substance.

3.3 The Core Argument

The hierarchical metaphysical framework just presented provides nearly all the materials needed in order to formulate Descartes's basic argument for God's existence in *Meditation III*, which we shall call his "core" argument. Indeed, the only further element we need to add is a general principle concerning the *causes* of ideas. It should come as no surprise that Descartes uses such a principle, for as previously noted, his argument is essentially a causal one, which reasons from the idea of God to God himself as the only possible cause of this idea. The principle that he uses is:

The cause of an idea must have as much formal reality as the idea contains objective reality.

It is crucial to understand that this principle does *not* say that an idea must have as much objective reality as it has formal reality: it does not compare ideas with ideas. It compares the degree of formal reality had by *the cause that produced an idea* with the degree of objective reality contained in that idea. Thus the principle says two different things: (1) an idea requires a cause, and (2) this cause must be an adequate one. In other words, not only must an idea have a cause, but the more objective

reality the idea has—the more reality it represents its object as having—the more formal reality its cause must have. Despite the abstract nature of Descartes's principle and the technical terminology he uses to express it, the principle is rather plausible. Descartes brings this out in his *Principles of Philosophy*, Part I, no. 17, by applying it to a concrete example:

[T]he greater the amount of objective perfection [= objective reality] they [= ideas] contain within themselves, the more perfect their cause must be. For example, if someone has within himself the idea of a highly intricate machine, it would be fair to ask what was the cause of his possession of the idea: did he somewhere see such a machine made by someone else; or did he make such a close study of mechanics, or is his own ingenuity so great, that he was able to think it up on his own, although he never saw it anywhere? All the intricacy which is contained in the idea merely objectively—as in a picture—must be contained in the cause, whatever kind of cause it turns out to be; and it must be contained not merely objectively or representatively, but in actual reality, either formally or eminently, at least in the case of the first and principal cause. (CSM I 198, SPW 165–6, AT VIIIA 11)

To see still better the plausibility of Descartes's principle, consider also the following elaboration of his example, by John Cottingham:

Recall Descartes' example of the “highly intricate machine”; and for the sake of simplicity let us follow Descartes' own comparison between ideas and pictures, and consider the case of a drawing or a picture rather than an idea. Suppose a five-year-old child produces a highly complicated design for a computer—a design which we know could only be produced by a highly skilled mathematician with a mental age vastly superior to the child's. The fact that the design is only a drawing and not an actual computer does not block the causal question: the representative or “objective” intricacy of the design still has to be accounted for. Of course, the child might simply have copied down the drawing from a book. But this simply pushes the argument one stage further back. We are, it seems, justified in asserting that somewhere along the line of causation there must be an actual entity or being that really does possess sufficient complexity to account for the

complexities which are to be found in the design. And what goes for the drawing goes equally for an idea: complex representational content requires a complex cause.¹¹

To arrive at Descartes's basic or "core" argument for God's existence, we need only relate the principle just presented to the metaphysical framework described in the previous section. Specifically, we need to ask whether the principle shows how any of our ideas—those of modes, finite substances, or infinite substance—are caused. Can it take us from knowledge of items on the left-hand side of our diagrams to knowledge of the existence of any of the items on the right-hand side? Let us consider first the ideas of modes or properties—those that contain the lowest degree of objective reality. These ideas may, in conformity with Descartes's principle, be caused by anything that has at least as much formal reality as they contain objective reality. So ideas of modes can be caused by modes, which have exactly as much formal reality as those ideas contain objective reality. But they can also be caused by finite substances, which have more formal reality than those ideas contain objective reality, or by God, who has even more formal reality. The mere fact that I have an idea of a mode, then, does not prove that any mode (other than that idea itself, which is a mode of me) actually exists, since this idea could be caused instead by a finite substance or by God. Consider, next, the ideas of finite substances. These ideas cannot, in conformity with Descartes's principle, be caused by modes. For a mode has less formal reality than the idea of a substance contains objective reality. So if a mode were to cause the idea of a finite substance, this would violate the principle that the cause of an idea must have as much formal reality as the idea contains objective reality. On the other hand, the idea of a finite substance can, in conformity with Descartes's principle, be caused in either of two different ways: by a finite substance, which has exactly as much formal reality as the idea contains objective reality, or by God, who has even more formal reality. So again, the mere fact that I have an idea of a finite substance cannot prove that such a substance exists, since this idea might be caused by an infinite substance (or by myself, since I am a finite thinking substance). So at this point, Descartes's principle has still not enabled him to know the existence of anything beyond

¹¹ John Cottingham, *Descartes*, pp. 52–53. For another helpful explanation of the same point, see Bernard Williams, *Descartes: The Project of Pure Enquiry*, pp. 138–39.

himself (and that knowledge, of course, stems only from the *cogito*, not from any principle about the causes of ideas). However, there still remains one idea to be considered—the idea of an infinite substance, or God. This is the idea of “a substance that is infinite, <eternal, immutable>, independent, supremely intelligent, supremely powerful, and which created both myself and everything else (if anything else there be) that exists” (CSM II 31, SPW 93–94, AT VII 45). How is this idea produced? It cannot, in conformity with Descartes’s principle, be caused by a mode; for a mode has far less formal reality than the idea of God, as just described, contains objective reality. But neither can it be caused by a finite substance; for such a substance still has less (infinitely less) formal reality than this infinitely rich idea of God contains objective reality. In fact, it is now obvious that there is only one way the idea of God could be caused. It could only be caused by God himself, for only God himself has as much formal reality as the idea of God contains objective reality. Thus, the idea of God differs from all other ideas. It is a uniquely privileged idea; for, alone among all ideas, the idea of God is such that from the mere fact that I have the idea, it follows that the *object* of that very idea—God himself—is also the *cause* of that idea and must, therefore, really exist.

This “core” argument for God’s existence can be formulated as follows:

1. The cause of an idea must have as much formal reality as the idea contains objective reality.
2. Only a perfect God has as much formal reality as my idea of God contains objective reality.
3. The cause of my idea of God is a perfect God (from propositions 1 and 2).
4. A perfect God really exists (from proposition 3).

We saw earlier that “objective reality” and “formal reality” are technical terms that Descartes uses in expressing the two key ideas involved in his argument: (1) that some ideas represent their objects as having more reality than other ideas represent their objects as having, and (2) that the degree of reality an idea represents its object as having depends on the degree of reality possessed by the object itself. Notice, then, that these two key ideas can be expressed without even using the terms “objective reality” and “formal reality,” since we in fact introduced those ideas before reexpressing them in those terms. Likewise, Descartes’s core

argument can be formulated without the terms “objective reality” and “formal reality.”¹² The formulation would go this way:

- 1'. The cause of an idea must have as much reality as the idea represents its object as having.
- 2'. Only a perfect God has as much reality as my idea of God represents God as having.
3. The cause of my idea of God is a perfect God (from propositions 1' and 2').
4. A perfect God really exists (from proposition 3).

So if you are having trouble grasping Descartes’s argument, this trouble should not stem from his technical terminology. Rather, it should stem from the substantive claims the argument makes—that there are different degrees of reality, that ideas represent their objects as having these different degrees of reality, that the idea of God represents God as having infinite reality, and that the only adequate cause of an idea’s representing its object as having a given degree of reality is a thing which actually has at least that degree of reality. We shall critically examine some of these claims in section 4.

3.4 The Central Argument of *Meditation III*: The Subargument, the Core Argument, and the Sequel

In *Meditation III*, Descartes does not formulate his core argument as concisely as we have just done. Rather, he embeds it within a longer argument that starts from certain very general principles about causality and ends with the vindication of his criterion of truth. His discussion can be broken down into three episodes: (1) an argument, which we will call “the subargument,” intended to establish the core argument’s first premise; (2) a survey of ideas that is designed to exhibit the uniquely privileged nature of the idea of God, and that culminates in Descartes’s discovery of the core argument’s second premise; and (3) an argument, which we will call “the sequel,” that goes from the core argument’s conclusion to the vindication of the “clarity and distinctness” criterion of truth. The subargument argument, core argument, and sequel constitute one continuous extended argument that is usually considered to be the

¹² The idea of restating Descartes’s argument without using his technical terms is due to Curley, *Descartes Against the Skeptics*, p. 126 (though Curley himself paraphrases out only the term “objective reality”).

central metaphysical argument of the *Meditations*. The purpose of this section is to present this extended argument.

The subargument begins this way:

Now it is manifest by the natural light that there must be at least as much <reality> in the efficient and total cause as in the effect of that cause. For where, I ask, could the effect get its reality from, if not from the cause? And how could the cause give it to the effect unless it possessed it? It follows from this both that something cannot arise from nothing, and also that what is more perfect—that is, contains in itself more reality—cannot arise from what is less perfect. (CSM II 28, SPW 91, AT VII 40–41)

Descartes's reasoning here starts from a premise, that can be extracted from his two rhetorical questions, and which we shall call the “precontainment principle.” It may be put this way:

(1) A cause must precontain the reality of its effect.

An alternative way to read the passage is to interpret Descartes's two rhetorical questions as containing the following argument for (1):

- (-1) An effect can get its reality only from its cause.
- (0) A cause can give its reality to its effect only if this cause possesses that reality.

∴ (1) A cause must precontain the reality of its effect.

In fact, this argument is maximally faithful to the text. However, it is not stronger than the version that starts with (1) as a basic premise. For (-1) means either that

- (-1a) An effect must *take* its reality from its cause—its reality must be *transferred* to it from its cause

or

- (-1b) An effect can be produced only by its cause.

But if (-1) means (-1a), then (-1) comes too close to (1) to avoid begging the question. On the other hand, if (-1) means (-1b), then (0) must be replaced by

- (0*) A cause can produce an effect only by precontaining the reality of that effect

which again comes so close to (1) that it begs the question.

Another reason not to favor the alternative reading over the one that starts with (1) is that there are several passages where Descartes says that (1) is the true starting-point of his argument. Thus, in one of his letters, he writes: “I proved this [i.e., that a mind could not have the idea of a perfect God unless such a God really existed] from the principle that there can be nothing in an effect which is not previously present in the cause” (to Hyperaspistes, August 1641, CSMK 192, AT III 428). That he regards the precontainment principle as absolutely basic is also confirmed by his reply to one of Gassendi’s objections, where Descartes refers to “the axiom ‘There is nothing in the effect which did not previously exist in the cause’” (CSM II 252, AT VII 366; my emphasis), and by a remark in his reply to the Second Set of *Objections*, where he gives, as the reason why “nothing cannot be the cause of a thing,” that “such a cause would not contain the same features as are found in the effect” (CSM II 97, AT VII 135). Of course, it can hardly be said that the meaning of the precontainment principle is clear, and later we shall have to ask what it means. Our present point is simply that the principle may fairly be taken as the opening premise of Descartes’s argument.

Descartes’s second premise, which he leaves unstated in the above passage, is

- (2) There are degrees of reality.

As we have already seen, this idea, which can be traced back at least to Plato, is essential to Descartes’s argument.

From these two premises, Descartes says that two things follow (“It follows both that . . . and that . . .”). First, he says it follows that

- (3) Something cannot come from nothing.

This claim, which might be called the “nothing-comes-from-nothing principle,” can be traced back to ancient Greek philosophy and in medieval

times was expressed in the Latin dictum *Ex nihilo, nihil fit* ("From nothing, nothing comes"). It means that something cannot be *caused* to exist or occur by nothing(ness) or nonbeing. Descartes derives it from (1). His reasoning is, presumably, that since a cause must precontain the reality of its effect, and since nothing(ness) cannot possibly contain anything, nothing(ness) cannot be the cause of anything. Second, Descartes says it follows that "what is more perfect—that is, contains in itself more reality—cannot arise from what is less perfect," in other words,

- (4) A cause must contain at least as much reality (perfection) as its effect.

This claim, which reasserts Descartes's paragraph's opening statement that "there must be at least as much <reality> in the efficient and total cause as in the effect of that cause," is a general principle that says that a cause must be *adequate* to its effect. We may call it, following John Cottingham, the "*Causal Adequacy Principle*".¹³ Descartes derives it from (1) and (2). His reasoning, presumably, is that since the cause must precontain the reality of its effect and given that the cause and the effect each have a certain degree of reality, the effect cannot have more reality than the cause; for if it did, then its "surplus" reality could not have been precontained in the cause.

The Causal Adequacy Principle was a commonplace of Scholastic philosophy. Descartes's innovation was to apply this principle to ideas, and in particular, to the "objective reality" or representational content of ideas. He does this in the continuation of the passage quoted above:

And this is transparently true not only in the case of effects which possess <what the philosophers call> actual or formal reality, but also in the case of ideas, where one is considering only <what they call> objective reality. A stone, for example, which previously did not exist, cannot begin to exist unless it is produced by something which contains, either formally or eminently everything to be found in the stone. . . . But it is also true that the idea . . . of a stone, cannot exist in me unless it is put there by some cause which contains at least as much reality as I conceive to be . . . in the stone. . . . [I]n order for a given idea to contain

¹³ Cottingham, *Descartes*, p. 49.

such and such objective reality, it must surely derive it from some cause which contains at least as much formal reality as there is objective reality in the idea. (CSM II 28–29, SPW 91, AT VII 41)

Descartes's language here is very compact. In the sentence about the stone, he is asserting two different things rather than only one. First, he is asserting that a stone cannot begin to exist “unless it is produced by something”—unless it has some cause. He then says the same thing about the *idea* of the stone: it, too, “cannot exist in me unless it is put there by some cause.” Second, he is asserting that the causes of the stone and of the idea of the stone must be adequate ones: they must be causes “which contain” enough formal reality to produce the effect. He seems not to distinguish these two points; his sentence structure fuses them together. Yet, as Bernard Williams has insightfully pointed out, there are two importantly different points here.¹⁴ One is the Causal Adequacy Principle, calling for a certain type of cause—an adequate one. The other is a principle that simply calls for a cause, period. So, in addition to the Causal Adequacy Principle, Descartes is here also invoking the principle that

(5) Everything must have a cause.

This principle, which is commonly called “the causal principle” or (following David Hume) “the causal maxim,” is different from—in a sense more basic than—the Causal Adequacy Principle; for the latter only says that *if* something has a cause, then that cause must have as much reality as its effect; it does not say *whether* a thing must have a cause. By contrast, the causal maxim asserts the universal need for a cause. (Hume actually formulates the maxim as “whatever has a *beginning* [emphasis added] of existence must have a cause of existence.” This has the advantage of not entailing that God must have a cause, since God is supposed to exist eternally. But some philosophers, including Descartes, would say that God is his own cause; not because he existed before he existed, which would be absurd, but because he exists necessarily, in virtue of his own nature.)

We already know the basis of the Causal Adequacy Principle—that is, of (4). It is derived from (1) and (2). But what is the basis for (5)? One possibility would be to say, as have some philosophers, that (5) needs no defense—that it is simply self-evident that everything must have a cause.

¹⁴ Williams, *Descartes: the Project of Pure Enquiry*, p. 141.

But Descartes does not treat (5) as self-evident, at least not in this passage; for he presents the need for a cause of the stone and the need for a cause of the idea of the stone as specific illustrations of, or as applications of, (3) and/or (4). So (5)—the general principle calling for those causes—is derived from one or both of those propositions. But (5) does not follow from (4): (4) calls for an adequate cause in cases where we admit that there is a cause but says nothing about whether there needs to be a cause in the first place. It seems, then, that Descartes takes (5) to follow from (3), or perhaps simply to be equivalent to (3). His reasoning seems to be that since something cannot come from nothing, everything must have a cause; he seems to derive the causal maxim from the nothing-comes-from-nothing principle. In the next section, we will see that there is a hidden difficulty (first spotted by Hume) in this reasoning.

Before continuing with the argument, we should address a delicate terminological matter. In the passage we have just discussed and in others that we shall encounter, Descartes uses the adverbial locution that the cause must “formally or eminently” contain all the reality of the effect. An examination of the relevant texts shows that for Descartes, whenever a cause formally or eminently contains all the reality of its effect, it has at least as much formal reality as its effect contains either formal or objective reality, and that is really the only point one needs to know to follow Descartes’s argument. But for readers who are interested (others can skip the next two paragraphs without losing track of the argument), we shall here offer an explanation of Descartes’s terminology.

There are three passages in the *Meditations* that are especially relevant to the question of what “formally or eminently contain” means. Here they are:

[A] A stone, for example, which previously did not exist, cannot begin to exist unless it is produced by something which contains, either formally or eminently everything to be found in the stone; similarly, heat cannot be produced by an object which previously was not hot, except by something of at least the same order <degree or kind> of perfection as heat, and so on. (CSM II 28, AT VII 41, SPW 91)

[B] As for all the other elements which make up my ideas of corporeal things, namely extension, shape, position, and movement, these are not formally contained in me, since I am nothing

but a thinking thing; but since they are merely modes of a substance, and I am a substance, it seems possible that they are contained in me eminently. (CSM II 31, AT VII 45, SPW 93)

[C] So the only alternative is that it [i.e., the cause of my ideas of sensible objects] is that it is in another substance distinct from me—a substance which contains either formally or eminently all the reality which exists objectively in the ideas produced by this faculty. . . . This substance is either a body, that is, a corporeal nature, in which case it will contain formally <and in fact> everything which is to be found objectively <or representatively> in the ideas; or else it is God; or some creature more noble than body, in which case it will contain eminently whatever is to be found in the ideas. (CSM II 55, AT VII 79, SPW 115)

These passages show that Descartes's distinction between formal and eminent containment does not introduce some third type of reality, "eminent reality," in addition to formal and objective reality; rather, as the adverbial locution "formally or eminently contain" implies, the distinction refers to the *way* in which reality, whether formal or objective, is contained in anything. Passages A and C show that one object can formally or eminently contain the reality contained in another object, such as a stone or a quantity of heat; and also that an object can formally or eminently contain the objective reality of an idea. Passage C shows that at least one way in which an object can eminently contain the formal reality of another object, or the objective reality of an idea, is by containing *more* (or a higher degree of) formal reality than the other object contains formal reality, or *more* formal reality than the idea contains objective reality. This point accords with John Cottingham's translator's note to the paragraph where Descartes expounds the subargument, in which Cottingham writes that "in scholastic terminology, to possess a property 'formally' is to possess it literally, in accordance with its definition; to possess it 'eminently' is to possess it in some higher form" (CSM II 28, note 2; SPW 91, note 1). Passage B strongly suggests, and Cottingham's note implies, a special requirement for it to be true that one thing, X, formally contains either the formal reality of another object Y or the objective reality of an idea I; namely, that the formal reality of X be not only equal to but also *of the same kind* as the formal reality of Y or the objective reality of I, where "kind" means either (a) mental or (b) physical. The reason why B strongly suggests this requirement is that Descartes gives, as

the reason why I do not formally contain the “other elements which make up my ideas of corporeal things, namely extension, shape, position, and movement” is that “I am nothing but a thinking thing.” So, the reason why I do not formally contain the reality of these elements but contain that reality only eminently is not, contrary to what one might think, merely that I contain *more* formal reality than they contain objective reality (or possibly formal reality: it isn’t clear whether “elements which make up my ideas of corporeal things” refers directly to the properties of those things, or to the representational content of the ideas of those properties, or to both of these), but rather that as a *thinking* thing I contain a different *kind* of reality than they do, by virtue of their involving extension and its modes. The requirement that formal containment involve sameness of ontological kind also seems to be implied by passage C, where Descartes says that only corporeal (i.e., bodily) nature formally contains the objective reality contained in ideas of bodies, and the gloss of “formally” as “in fact” (added in the French translation as *en effet*) reinforces this requirement as well.

Against the background of these textual points, we can see the purpose of Descartes’s distinction between formal and eminent containment. It is to allow for four different ways in which a cause can contain at least as much reality as its effect—four ways in which the Causal Adequacy Principle can be satisfied. First, the cause may contain exactly the same degree and the same kind (i.e., physical or else mental) of formal reality as the effect contains either objective reality (if it is an idea) or formal reality. In that case, Descartes says that the cause *formally contains* all the reality that the effect contains, or formally contains “everything to be found in” the effect. Second, the cause can contain more than but the same kind of formal reality as the effect contains objective reality or formal reality. Third, the cause can contain at least as much as but a different kind of formal reality than the effect contains formal or objective reality. Fourth, the cause can contain both more than and a different kind of formal reality than the effect contains objective or formal reality. In the second, third, and fourth cases, Descartes says that the cause *eminently contains* all the reality that the effect contains. A simpler way to put all of this is to say that a cause formally contains the reality of its effect when it has exactly the same degree and kind of formal reality as its effect has objective or formal reality, and that a cause eminently contains the reality of its effect in all other cases where it contains at least as much formal reality as its effect contains formal or objective reality, that is, in

all cases where it does not formally contain that reality. The most important point for Descartes's argument, as previously noted, is that whenever a cause either formally or eminently contains contains the objective reality of an idea, the first premise of his core argument is true: the cause of the idea contains at least as much formal reality as the idea contains objective reality.

To return to the argument—assuming that (5) is established, Descartes now applies (5) to the “objective reality” of ideas. The objective reality (representational or informational content) of an idea of a stone (i.e., the fact that this idea represents a stone as having a certain degree of reality) must have a cause, no less than the fact that a stone exists. Descartes's next step, then, is to derive from (5) the following:

- (6) The objective reality of an idea must have a cause.

To arrive at the core argument's first premise, Descartes uses one further premise. That premise is contained in the following passage:

And although the reality which I am considering in my ideas is merely objective reality, I must not on that account suppose that the same reality need not exist formally in the causes of my ideas, but that it is enough for it to be present in them objectively. For just as the objective mode of being belongs to ideas by their very nature, so the formal mode of being belongs to the causes of ideas—or at least to the first and most important ones—by their nature. And although one idea may perhaps originate from another, there cannot be an infinite regress here; eventually one must reach a primary idea, the cause of which will be like an archetype which contains formally <and in fact> all the reality <or perfection> which is present only objectively <or representatively> in the idea. (CSM II 29, SPW 91–92, AT VII 41–42)

Here Descartes admits that the objective reality of an idea is sometimes borrowed or derived from the objective reality of one or more other ideas. A modern example might be that an engineer's idea of a rocket engine (one that carries its own “air” supply as well as its fuel) could be derived from the engineer's idea of a jet engine (one that takes its air supply from outside): perhaps the engineer constructed his idea of a rocket engine

from elements in his idea of a jet engine. But Descartes would insist that such an explanation of how the engineer's idea of a rocket engine originated is incomplete. For now we need an explanation of how the engineer acquired the idea of a jet engine. In the end, Descartes would say, this explanation cannot be just the fact that the engineer constructed it from yet another idea. Rather, it must be either that the engineer has such a brilliant and inventive mind that he created the idea of a jet engine or the idea(s) from which it is constructed, or else that he got that idea from observing a jet engine itself or from something having the same as or a higher level of complexity than a jet engine. In other words, ultimately, the cause of an idea's representational content cannot be just the representational content of another idea. It must be some nonrepresentational fact about the world. Representational content is, in the end, parasitic on nonrepresentational facts.

Perhaps the following analogy can help to bring out Descartes's point. Imagine a mirror image. This mirror image could itself be a reflection of an image reflected by a second mirror. And the image in the second mirror could be the reflection of an image reflected by a third mirror. Indeed, the mirrors could be so arranged that there might be a very long series of mirrors, each reflecting a mirror image reflected by its predecessor in the series. But this series could not go on infinitely. It must terminate in a mirror image which is not itself the image of an image, but which is caused by something other than a mirror image. The same holds for the representational content of an idea: it may be derived from another idea's content—and that from yet another idea's content. But such a series cannot continue infinitely; in the long run it must terminate in a content that is caused by something other than an idea's content. In Descartes's terminology, then, the objective reality of an idea must ultimately be caused by the "formal reality" of something, not just by the objective reality of another idea.

Descartes's premise, then, could be put this way:

Although this cause may be the objective reality of another idea or ideas, ultimately (or in the long run) it must be the formal reality of something.

To simplify matters a little, however, let us formulate the premise as:

(7) This cause must be the formal reality of something.

We have now assembled all the materials Descartes offers in support of the core argument's first premise: (4) says that a cause must contain as much reality as its effect, (6) says that the objective reality of an idea must have a cause, and (7) says that this cause must be the formal reality of something. From these three statements, Descartes derives the premise

- (8) The cause of an idea must have as much formal reality as the idea contains objective reality.

The next episode of the *Meditation* is not, strictly speaking, part of the extended argument (subargument, core argument, and sequel). Rather, its function is to bring out the unique nature of the idea of God, thereby preparing the way for the core argument's second premise. To some extent, we have already anticipated this material in our explanation of the core argument in the previous section. But it is worth seeing briefly how Descartes himself presents the matter. He begins dramatically, by announcing an implication of (8) and of the weakness of his previous reasons for thinking that some of his ideas are caused by things outside himself:

If the objective reality of any of my ideas turns out to be so great that I am sure the same reality does not reside in me, either formally or eminently, and hence that I myself cannot be its cause, then it will necessarily follow that I am not alone in the world, but that some other thing which is the cause of this idea also exists. But if no such idea is to be found in me, I shall have no argument to convince me of the existence of anything apart from myself. For despite a most careful and comprehensive survey, this is the only argument I have so far been able to find. (CSM II 29, SPW 92, AT VII 42)

In other words, (8) now offers the only possible way of escaping from *solipsism*, that is, from the extraordinary view that only I and my own thoughts exist.

Descartes now proceeds, in effect, by raising the following causal question: is the objective reality of any of my ideas such that the idea's *object* (what the idea is of or about) must also be the idea's *cause*? If the answer is yes, then this will show that the object of the idea cannot be nonexistent or merely fictitious, but must really exist, in order to cause the idea.

So Descartes makes an inventory of his ideas, with a view to answering his causal question. He finds that, in addition to his idea of himself as a finite, thinking substance, he has ideas of

- (a) God,
- (b) inanimate physical objects,
- (c) other humans, animals, and angels.

My ideas of other humans, animals and angels, says Descartes, could easily be constructed by combining elements taken from my ideas of myself, of inanimate physical objects, and of God even if no other humans, animals or angels existed. So my having ideas in class (c) does not show that other humans, animals, or angels exist. My ideas of inanimate physical objects' properties could originate entirely from myself—if only because they are ideas of modes, which contain less objective reality than the formal reality I myself contain as a finite substance. My ideas of those objects themselves could also originate from myself, since they are ideas of finite substances and I, as a finite thinking substance, possess sufficient formal reality to cause such ideas. So my having ideas in class (b) does not show that physical objects really exist. So, says Descartes,

[T]here remains only the idea of God; and I must consider whether there is anything in the idea which could not have originated from myself. By the word “God” I understand a substance that is infinite, <eternal, immutable>, independent, supremely intelligent, supremely powerful, and which created both myself and everything else (if anything else there be) that exists. All these attributes are such that, the more carefully I concentrate on them the less possible it seems that the idea I have of them could have originated from me alone. So from what has been said it must be concluded that God necessarily exists. (CSM II 31, SPW 93–94, AT VII 45)

Here Descartes finds, in the first place, that the idea of God contains more objective reality than Descartes possesses formal reality. So, his having this idea does show, finally, that something other (and greater) than himself exists—that he is “not alone in the world,” that solipsism is false. But further, Descartes finds that only one entity possesses enough

formal reality to cause this great idea, namely, God himself. So God must really exist.

Descartes's survey of ideas, then, has culminated in the completion of his core argument; for the survey has led to his discovering the core argument's second premise,

- (9) Only a perfect God has as much formal reality as my idea of God contains objective reality.

But from (8) and (9), there follows

- (10) The cause of my idea of God is a perfect God.

Finally, it obviously follows from (10) that

- (11) A perfect God exists.¹⁵

As important as is this conclusion, it is not yet the final conclusion of Descartes's extended argument; for remember what Descartes's project in *Meditation III* was—to vindicate his criterion of truth. To achieve this purpose, Descartes declared that he needed to show both that (a) God exists and (b) God is not a deceiver. We have yet to see how he establishes (b) and exactly how (b) is supposed to vindicate the “clarity and distinctness” criterion of truth. To see this, we must look at the third and final stage of Descartes's extended argument—the part we are calling “the sequel.”

The sequel's key premise is this:

- (12) To deceive is an imperfection.

From this premise, together with (11), Descartes concludes,

- (13) God is not a deceiver.

¹⁵ It is also possible to read the episode just discussed as providing an alternative proof of this conclusion, as follows:

- i. The object of an idea must also be its cause just in case only that object has enough formal reality to cause this idea.
- ii. Only a perfect God has enough formal reality to cause my idea of God.
- iii. The cause of my idea of God is a perfect God. [from (i) and (ii)]
- iv. A perfect God exists. [from (iii)]

Descartes presents this brief piece of reasoning in a number of places, saying that its premise is known “by the natural light.” For example, near the end of *Meditation III*, he writes:

By “God” I mean . . . the possessor of all . . . perfections . . . who is subject to no defects whatsoever. It is clear enough from this that he cannot be a deceiver, since it is manifest by the natural light that all fraud and deception depend on some defect. (CSM II 35, SPW 98, AT VII 52)

Again, at the beginning of *Meditation IV*, he says:

To begin with, I recognize that it is impossible that God should ever deceive me. For in every case of trickery or deception some imperfection is to be found; and although the ability to deceive appears to be an indication of cleverness or power, the will to deceive is undoubtedly evidence of malice or weakness, and so cannot apply to God. (CSM II 37, SPW 99, AT VII 53)

Finally, Descartes maintains, (13) vindicates the “clarity and distinctness” criterion of truth. This can be seen in the *Synopsis* of the *Meditations*, where he announces.

In the Fourth Meditation it is proved that everything that we clearly and distinctly perceive is true . . . (CSM II 11, SPW 75, AT VII 15)

The passage in *Meditation IV* to which he is referring is this:

[E]very clear and distinct perception is undoubtedly something, and hence cannot come from nothing, but must necessarily have God for its author. Its author, I say, is God, who is supremely perfect, and who cannot be a deceiver on pain of contradiction; hence the perception is undoubtedly true. (CSM II 43, SPW 105, AT VII 62)

In *Meditation V*, he makes the point this way:

Now, however, I have perceived that God exists, and at the same time I have understood that everything depends on him, and that he is no deceiver; and I have drawn the conclusion that everything which I clearly and distinctly perceive is of necessity true. (CSM II 48, SPW 109, AT VII 70)

These passages show that Descartes believes he can go from “God is not a deceiver” to “whatever I clearly and distinctly perceive is true.” But they do not show exactly how the one proposition is supposed to lead to the other. So let us try to fill in the missing step(s) in the reasoning. The key premise that we need is this:

- (14) If my clear and distinct perceptions could be false, then God would be a deceiver.

This premise rests on the “assent-compelling” nature of clear and distinct perceptions. Recall that Descartes says that while one is having a clear and distinct perception, one *cannot* doubt it: one is compelled to assent to it. Thus, if one of my clear and distinct perceptions could nevertheless be false, then I would be making an error that I was powerless to correct: I would be irremediably deceived. And since God created me and my cognitive equipment, this would mean that he was a deceiver.

With the help of (14), Descartes can at last derive his final conclusion; for it follows from (13) and (14) that

- (15) My clear and distinct perceptions cannot be false, that is, whatever I clearly and distinctly perceive is true.

This completes Descartes’s extended argument. Descartes’s criterion of truth, originally extracted from the *cogito*, is now finally vindicated by God’s veracity (truthfulness).

Since the extended argument is quite lengthy, it may be helpful to summarize it, by listing all of its steps in one place. Here, then, is the entire argument. Steps (1)–(8) are the “subargument”; steps (8)–(11) are the “core argument”; steps (11)–(15) are the “sequel.”

1. A cause must precontain the reality of its effect. (premise).
2. There are degrees of reality (premise).
3. Something cannot come from nothing (from step 1).

4. A cause must contain at least as much reality as its effect (from steps 1 and 2).
5. Everything must have a cause (from step 3).
6. The objective reality of an idea must have a cause (from step 5).
7. This cause must be the formal reality of something (premise).
8. The cause of an idea must have as much formal reality as the idea contains objective reality (from steps 4, 6, and 7).
9. Only a perfect God has as much formal reality as my idea of God contains objective reality (premise).
10. The cause of my idea of God is a perfect God (from steps 8 and 9).
11. A perfect God exists (from step 10).
12. To deceive is an imperfection (premise).
13. God is not a deceiver (from steps 11 and 12).
14. If my clear and distinct perceptions could be false, then God would be a deceiver (premise).
15. My clear and distinct perceptions cannot be false, that is, whatever I clearly and distinctly perceive is true (from steps 13 and 14).

In concluding this section, we should note that Descartes thought that his proof that a nondeceiving God exists raises a problem of its own. This problem, which is the main topic of *Meditation IV*, is the following: How does it happen that I sometimes make errors (hold false beliefs), if God is not a deceiver? By proving that God is not a deceiver, hasn't Descartes proved "too much," that is, proved that error is impossible? The gist of Descartes's answer to this question is that error arises from misusing my free will, by adopting beliefs even about matters that my intellect does *not* clearly and distinctly perceive. So long as I restrict myself to matters that I do clearly and distinctly perceive, error is impossible (CSM II 40–1, SPW 102–3, AT VII 58–59). In the following chapter, we shall analyze *Meditation IV* and consider more closely Descartes's theory of error. But in the rest of the present chapter, we shall offer an assessment of Descartes's extended argument.

4. Criticisms of Descartes's Central Argument in *Meditation III*

Descartes's central argument in *Meditation III* raises many issues and can be criticized in a number of different ways. We shall consider three possible lines of criticism: one focusing on the subargument, a second on

Descartes's overall strategy, and a third on the core argument's second premise.

4.1 The Subargument

Our critical discussion of the subargument will focus on its first five steps (see p. 49). We shall discuss three difficulties in those five steps.

4.1.1 The Precontainment Principle

The first difficulty is that premise (1), the precontainment principle, involves a highly problematic conception of causation (cause-and-effect). Exactly what does it mean to say that a cause must precontain the reality of its effect, or that "there can be nothing in an effect which is not previously present in the cause" (letter to Hyperaspistes, August 1641, CSMK 192, AT III 428)? In this subsection, we shall examine some possible answers to this question.

The simplest interpretation of the precontainment principle would be that the cause must actually precontain its effect, much as a fetus is pre-contained in the womb. In an illuminating article, the distinguished historian of ideas Arthur O. Lovejoy (1873–1962) discusses what he calls the "preformationist assumption about causality." He writes,

That "there cannot be more in the effect than there is in the cause" is one of the propositions that men have been readiest to accept as axiomatic; a cause, it has been supposed, does not "account for" its effect, unless the effect is a thing which the eye of reason could somehow discern in the cause, upon a sufficiently thorough analysis.¹⁶

Lovejoy goes on to show that this "preformationist assumption" can be traced back to antiquity, was pervasive both in medieval and in seventeenth- and eighteenth-century philosophy, and continued to be put forward by some twentieth-century thinkers.

If the precontainment principle is interpreted in this way, however, then it is open to a serious objection: it seems to be simply false. This was

¹⁶ Arthur O. Lovejoy, "The Meanings of 'Emergence' and its Modes," p. 286.

powerfully argued by David Hume (1711–1776), whose views about causation remain enormously influential today. Hume wrote:

The mind can never possibly find the effect in the supposed cause, by the most accurate scrutiny and examination. For the effect is totally different from the cause, and consequently can never be discovered in it. Motion in the second Billiard-ball is a quite distinct event from motion in the first; nor is there anything in the one to suggest the smallest hint of the other . . . [E]very effect is a distinct event from its cause. It could not, therefore, be discovered in the cause.¹⁷

Hume's basic point is that if the effect were contained in the cause, then it would be possible for us to find or discern the effect by carefully examining the cause. But we cannot do this; rather, we must wait for experience to show us what effect will follow from any given cause.

In making this point, Hume explicitly refers to the cause and the effect as being each an *event*. This reflects an important insight of Hume's, namely, that the true members of a cause-effect relationship are events, rather than objects. Often, our ordinary speech masks this fact. For example, we say that “the rock broke the window.” Here it almost sounds as if the cause is one object (the rock), and the effect another (the broken window). But, of course, what really happened is that the rock's hitting the window caused the window's breaking. Now *the rock's hitting the window* and *the window's breaking* are not objects or things; they are events or occurrences. Once we understand this point, the idea of the cause's “containing” the effect immediately looks suspect. For it makes little if any sense to say that the rock's hitting the window “contained” the window's breaking: it certainly did not do so in the literal sense in which, for example, a box of chocolates contains the chocolates or the chocolates contain their caramel fillings. Of course, there are cases of causality that seem to fit the “preformationist assumption” much better, such as the case of conception and birth. But in such a case, it is true only in a general, rough sense that the mother “caused” the baby. No biologist studying the process of reproduction would describe what happened in such an inexact way. Rather, what really happened is that a

¹⁷ David Hume, *An Enquiry Concerning Human Understanding*, sec. 4, part 1, para. 9 and 11.

complex sequence of events involving the mother caused the event of the baby's birth. And again, it makes little if any sense to say that each of those events "contained" their effects. It would seem, then, that if Descartes's precontainment premise is interpreted to mean that the cause must literally precontain its effect, then the points made by Hume show the premise to be erroneous.

In all probability, however, Descartes's principle should not be interpreted so literally; for after all, his subargument's premise 1 says that the cause must precontain *the reality of* the effect, not that it must precontain the effect itself. So perhaps his premise means that in cases where the cause is a change in or an act of some object(s) (loosely called "the cause") and the effect is the coming into existence of another object(s) (loosely called "the effect"), the latter object(s) cannot possess any property that was not also possessed by the former. John Cottingham, in his *Descartes*, interprets Descartes's premise along these lines, calling it the "heirloom" view of causation:

To say that whatever produces a stone must itself have all the features found in the stone seems to imply a kind of "heirloom" view of causation—that the only way an effect can have come to possess some property is by inheriting it, heirloom fashion, from its causes.¹⁸

As Cottingham goes on to argue, however, "the heirloom principle" is highly questionable, for it seems vulnerable to counterexamples. Thus, the author(s) of the Second set of *Objections* to the *Meditations* wrote to Descartes:

You say . . . that an effect cannot possess any degree of reality or perfection that was not previously contained in the cause. But we see that flies and other animals, and also plants, are produced from sun and rain and earth, which lack life. . . . [H]ence it does happen that an effect may derive from its cause some reality which is nevertheless not present in the cause. (CSM II 88, AT VII 123)

To this Descartes replied that if animals have perfections that are not present in the sun, rain and earth, then this only shows that the sun, rain

¹⁸ Cottingham, *Descartes*, pp. 50–51.

and earth are not the total causes of animal life (CSM II 96, AT VII 134). As Cottingham points out, however, this reply commits Descartes to the view that there can never be genuinely emergent properties—that is, properties that were never possessed, in any previous state of the world, by the things whose operations caused those properties to occur. But this conflicts with the view, held by evolutionary biologists, that consciousness evolved from nonconscious forces and elements. Of course, even today, there are people who deny that consciousness evolved from inanimate elements. But the damaging point, so far as Descartes's argument is concerned, is that he could not even allow that there is a legitimate controversy here; for according to him, the precontainment principle is supposed to be obvious to any rational mind, quite apart from any empirical evidence provided by sciences like chemistry and biology.

Perhaps, however, the precontainment principle can be interpreted even more charitably. In his book *The Miracle of Theism*, John Mackie writes:

Though we ordinarily admit that great effects can be brought about by very small causes, these can only be partial causes, not the whole cause of the great effects. Large trees can grow from small seeds, but only by taking in a lot of nourishment as they grow. Trivial accidents can precipitate revolutions, but only where there are great repressed forces waiting to be triggered or released. And so on. We commonly assume that there are conservation principles—the conservation of mass, or of energy, or of the sum of the two, or of momentum—which operate as constraints on possible processes of causation or production or growth. Descartes's dictum that there must be as much reality in the total cause as in the effect can be understood as an attempt to capture the general form of which such specific conservation principles are instantiations.¹⁹

Mackie's remarks are directly addressed to (4), the causal adequacy principle, rather than to (1), the precontainment principle, from which Descartes derives (4). But they can be seen as offering a possible defense of (1) itself, no less than of (4). This defense would be that (1), like (4), is a very general principle of conservation, sanctioned by science. The basic

¹⁹ John Mackie, *The Miracle of Theism*, p. 35.

idea would that if the reality of an effect were not somehow precontained in the cause of that effect, this would violate a general conservation principle.

It is not necessary for us to try to spell out this idea more carefully; for if Descartes's premise is interpreted as a general conservation principle, then it cannot serve his purposes, since, as Mackie goes on to point out,

[Such a principle] is not known or knowable *a priori*. We have no rational guarantee, apart from experience, and apart from scientific theories developed from and confirmed by such experience, that it will hold . . . It cannot do the work that Descartes intends it to do in an absolutely secure rebuilding of human knowledge, since it is itself supported only by a wide range of interpreted observation, and its precise scope and implications are uncertain.²⁰

The key point here is that the premises of Descartes's argument must be knowable without dependence on what could be learned only from observing the operations of the physical world; for the argument is supposed to show that God's existence can be known without any use of the senses and even if the existence of the physical world is still in doubt. But as Mackie indicates, (1) and/or (4), construed as general conservation principles, can be known only by widely based observations of the physical world. So they cannot, construed in that way, play the role demanded of them by Descartes's argument.

Finally, it can be argued that even if one were to assume knowledge of the physical world, conservation principles of the kind Mackie has in mind cannot legitimately be used to argue for God's existence. For such principles, being ultimately based on our observations of physical things' operations and interactions, concern what we might call the "internal structure" of the physical universe—the regular qualitative—and, especially, quantitative—relationships among different spatial and temporal parts of the universe. But, as Immanuel Kant argued in his *Critique of Pure Reason* (1781), we are not justified in assuming that these principles therefore also provide reliable information about how the universe *as a whole* is related to a source that is supposed to have created it.

²⁰ Mackie, *The Miracle of Theism*, pp. 35–36.

We have considered three possible interpretations of Descartes's pre-containment principle. It can be interpreted as a statement of what Lovejoy calls the "preformationist assumption about causality," as a statement of what Cottingham calls the "heirloom view of causation," and as a statement of what Mackie calls "the general form" of "specific conservation principles." We have seen that on the first two interpretations, the principle is highly questionable and that on the third interpretation, it cannot serve Descartes's purposes. This result, of course, does not show that the conclusion of the subargument, namely (8) (which is the core argument's first premise, and which it is the whole purpose of the subargument to justify), is *false*. But it calls into serious question Descartes's attempt to justify or support that premise by appealing to the precontainment principle.

Before concluding this section, therefore, we should note that in one place, Descartes tries to establish his premise without relying on the precontainment principle. In his *Reply to the Second set of Objections*, Descartes says:

The fact that "there is nothing in the effect which was not previously present in the cause, either in a similar or in a higher form" is a primary notion which is as clear as any that we have; it is just the same as the common notion that "Nothing comes from nothing." For if we admit that there is something in the effect that was not previously present in the cause, we shall also have to admit that this something was produced by nothing. (CSM II 97, AT VII 135)

Here, Descartes seems to be defending a key premise needed to secure (8), namely (4), the Causal Adequacy Principle, which says that a cause must contain as much reality as its effect, without relying on (1), the pre-containment principle. The first three steps of his subargument would then be these:

- (2) There are degrees of reality.
- (3) Something cannot come from nothing.
- (4) A cause must contain at least as much reality as its effect
(from steps 2 and 3).

In the passage just cited, Descartes can be read as asserting that (4) follows from (2) and (3), because if an effect did contain more reality than

its cause, then it would have to get some of its reality from nothing, which violates (3). Since this way of supporting (4) makes no use at all of the precontainment principle, but only of the degrees of reality premise and the nothing-comes-from-nothing principle, it might seem to show that Descartes's subargument can proceed without using the precontainment principle. It must be said that in reading Descartes this way, we are bending over backward to be charitable; for in the very next sentence after the quoted passage, he defends (3)—the nothing-comes from nothing principle—by appealing back to the precontainment principle:

And the reason why nothing cannot be the cause of a thing is simply that such a cause would not contain the same features as are found in the effect. (CSM II 97, AT VII 135)

But this appeal to the precontainment principle is quite unnecessary, for there is a much better reason why nothing cannot be a cause of a thing: nothing—being just nothing, or nothingness, or nonbeing—obviously cannot cause anything. So despite Descartes's remark, this version of his subargument could proceed without relying on the precontainment principle.

Unfortunately, however, this alternative version of the subargument suffers from a fatal flaw; for, as James Van Cleve has shown, it subtly begs the question.²¹ For why should one accept Descartes's claim that if an effect e contains more reality than its cause c , then the "surplus" reality must have come from nothing? Why not maintain, instead, that this "surplus" reality all comes from the lesser cause, c ? Descartes would no doubt reply that the lesser cause c cannot be the cause of the greater effect e —that c "spent itself" in producing the "nonsurplus" part of e 's reality. But this assumes exactly the point to be proved, namely, that a cause must contain at least as much reality as its effect.

To conclude, it seems clear, in light of Van Cleve's criticism, that the alternative version of the subargument Descartes offers in his *Reply to the Second Set of Objections* is unsuccessful. Descartes must therefore rely on the version he gives in *Meditation III*, which is built on the highly problematic precontainment principle.

²¹ James Van Cleve, in personal correspondence and in "On a Little-noticed Fallacy in Descartes," manuscript.

4.1.2 Degrees of Reality

A second difficulty in the subargument concerns premise 2, that there are degrees of reality. Thomas Hobbes, author of the third Set of *Objections*, crisply challenged this premise:

Moreover, M. Descartes should consider afresh what “more reality” means. Does reality admit of more and less? Or does he think one thing can be more of a thing than another? If so, he should consider how this can be explained to us with that degree of clarity that every demonstration calls for, and which he has himself employed elsewhere. (CSM II 130, AT VII 185)

Descartes's reply was this:

I have also made it quite clear how reality admits of more and less. A substance is more of a thing than a mode; if there are real qualities or incomplete substances, they are things to a greater extent than modes, but to a lesser extent than complete substances; and, finally, if there is an infinite and independent substance, it is more of a thing than a finite and dependent substance. All this is completely self-evident. (CSM II 130, AT VII 185)

Here, Descartes seems to complicate his three-level hierarchy (infinite substance, finite substance, and modes) by adding an intermediate level between finite substances and modes: “real qualities or incomplete substances.” In his *Reply* to the fourth Set of *Objections*, he gives an example of an incomplete substance: a hand (CSM II 157, AT VII 222). It is a substance because it is a material thing like a rock or a stick but an incomplete one because its function cannot be understood apart from the whole body of which it forms a part. As Bernard Williams points out, however, these points are “basically unassimilated relics in Descartes’s metaphysics” of Aristotelian views that play no role within Descartes’s own scheme.²² And as Anthony Kenny points out, Descartes himself elsewhere strongly rejects the theory of “real accidents.”²³ So the real import of the above passage is that Descartes is reasserting, as “completely

²² Williams, *Descartes: The Project of Pure Enquiry*, p. 137.

²³ Anthony Kenny, *Descartes*, p. 133.

self-evident,” his three-level hierarchy. In the same vein, in his *Reply* to the second set of *Objections* (where he complies with a request that he present his whole system in the geometric format of definitions, postulates, axioms, propositions, and corollaries), he treats it as an axiom (a self-evident starting point) that

There are various degrees of reality or being: a substance has more reality than a mode; an infinite substance has more reality than a finite substance. Hence there is more objective reality in the idea of a substance than in the idea of a mode; and there is more objective reality in the idea of an infinite substance than in the idea of a finite substance. (CSM II 117, AT VII 165–166)

Descartes, then, seems to have considered his theory of degrees of reality so obvious as not to require any defense or explanation. Yet one can certainly question whether the basis of his hierarchy is clear. The standard reading of the hierarchy is that it turns on the idea of dependence and independence: if X can exist independently of Y but not vice-versa, then X has more (formal) reality than Y does. It can then be suggested that since finite substances depend on infinite substance, or God, for their existence but God does not depend on them (or on anything else) for his existence, infinite substance has more reality than finite substances and that since modes or properties cannot be “free-floating” but must exist in a substance, finite substances have more reality than modes or properties. This suggestion works well for the case of God and finite substances: the latter depend causally on God for their existence, but not vice-versa. Indeed, Descartes holds that finite substances depend causally on God in a very radical way: God not only creates them, but also preserves them in existence from moment to moment, or continually recreates them; for Descartes thinks that it is “evident by the natural light . . . to anyone who attentively considers the nature of time that the same power and action are needed to preserve anything at each individual moment as would be required to create that thing anew if it were not yet in existence” (CSM II 33, SPW 96, AT VII 49). But the suggestion does not work for finite substances and modes or properties. To see why, recall the account of the substance theory in chapter 2, section 6.1, according to which the minimum that can exist on its own is a substance *plus* a property: there can be no property without a substance, *and* there can be no substance without a property. Thus, there is a two-way, rather

than (as is sometimes thought) only a one-way, relation of dependence between substance and property. As Bishop Berkeley (1685–1753), the great Irish philosopher, put it, “It seems no less absurd to suppose a substance without accidents, than it is to suppose accidents without a substance.”²⁴ It remains unclear, then, why properties are supposed to have less reality than substances. Is it perhaps because a substance can undergo change without losing its identity, whereas a property cannot? Perhaps, but then Descartes’s hierarchy is based on two different notions of degrees of reality, rather than only one.²⁵ In any case, his theory of degrees of reality can hardly be said to be “clear and distinct,” or to be as unquestionable as the inference from “I am thinking” to “I exist.”

4.1.3 Justifying the Causal Maxim

A third problem with the subargument is that Descartes’s justification for (5), the causal maxim that everything must have a cause, is faulty. As we saw, Descartes evidently believed that (5) follows logically from (3), the nothing-comes-from-nothing principle, that something cannot come from nothing. Indeed, he may have thought, as did many philosophers before him, that (3) and (5) are logically equivalent, that is, that each statement follows logically from the other, so that they make the same assertion in two different ways. But there is a hidden mistake here. To expose this mistake, we need to see that (3) is ambiguous. It can mean

(3a) Nothing (= nothingness or nonbeing) cannot be a cause.

Or it can mean

(3b) Something cannot exist *without* a cause.

²⁴ George Berkeley, *A Treatise Concerning the Principles of Human Knowledge*, Part I, section 67.

²⁵ Ms. Chelsie Mack, a student at the College at Brockport, SUNY, wrote an insightful remark that may blunt the force of this criticism: “Finite substances such as the Meditator, or the lump of wax the Meditator mused about in [Meditation II], are ‘more real’ than modes or properties in the sense that finite substances have a degree of independence from their modes and properties. For instance, the wax’s shape, size, temperature and consistency (i.e., its properties) changed, but it continued to be the same piece of wax. Similarly, the Meditator has different thoughts and ideas, but remains the same ‘thinking thing.’”

The first philosopher to notice the ambiguity of (3) and to distinguish between (3a) and (3b) was David Hume, in a famous section of his *A Treatise of Human Nature* (1739).²⁶ Hume went on to point out, in effect, that even if (3a) is self-evident, this does not mean that (3b) is self-evident and to argue that (3b) is not self-evident. Furthermore, since (5) follows only from—and is indeed equivalent to—(3b) and does not follow from (3a), Hume concluded that (5) is not self-evident either.

Hume's distinction raises a problem for Descartes's subargument. Suppose that (3) means (3a). Then (3) follows from (1), because

(1) A cause must precontain the reality of its effect.

.: (3a) Nothing cannot be a cause.

is a valid argument, since nothing (nothingness, nonbeing) cannot possibly precontain anything. However, (5) does not then follow from (3), because

(3a) Nothing cannot be a cause.

.: (5) Everything must have a cause.

is an invalid argument, since the fact that nothing(ness) cannot serve as a cause does not prove that a thing could not just spring into existence *without* any cause whatsoever, including “nothing(ness)” or “nonbeing.” On the other hand, suppose that (3) means (3b). Then (5) does follow from (3), because

(3b) Something cannot exist without any cause.

.: (5) Everything must have a cause.

is valid (indeed, statements 3b and 5 are logically equivalent). However, (3) then does not follow from (1), because

(1) A cause must precontain the reality of its effect.

.: (3b) Something cannot exist without any cause.

²⁶ David Hume, *A Treatise of Human Nature*, book I, Part 3, sec. 3.

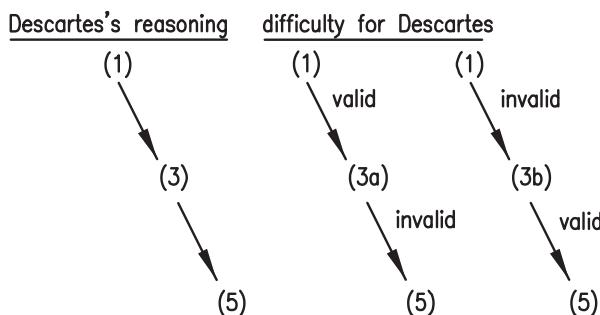


Figure 3.4

is an invalid argument. It is invalid because, as Bernard Williams has emphasized, (1) only covers cases where we admit that there is a cause; it says nothing at all about the possibility of something existing without any cause.²⁷

Using the statement numbers already assigned to represent the statements themselves, and an arrow to represent “therefore,” we can summarize the difficulty diagrammatically as shown in Figure 3.4. In brief: Descartes can go validly from (1) to (3) or from (3) to (5), but he cannot go validly from (1) to (3) to (5).

To conclude this section, we shall suggest two possible solutions to the difficulty just discussed. Each of these solutions is designed to put (8), which it is the whole purpose of the subargument to justify, in the strongest possible light. The first possible solution would be to treat (5)—and also (3b), since these are equivalent—as a basic premise or axiom. Using numbers and arrows the same way as in Figure 3.3 above, and using “+” to show that the statements linked by the “+” are intended to support the statement below them not individually but jointly, the entire subargument’s structure could then be diagrammed as shown in Figure 3.5.²⁸

The advantage of this solution is that the subargument would then not rely on the flawed method of justifying (5) just criticized. On the other hand, the argument would now rely on four basic premises (i.e., premises that are not themselves supported by arguments but used as

²⁷ Williams, *Descartes: The Project of Pure Enquiry*, p. 141.

²⁸ The method of diagramming arguments used here is explained, among other places, in Stephen Thomas, *Practical Reasoning in Natural Language*, pp. 57–73 and pp. 83–93.

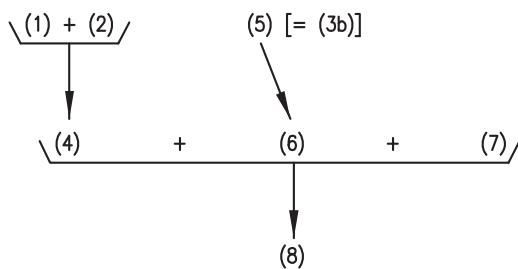


Figure 3.5

starting points of the argument): (1), (2), (5) = (3b), and (7). If the argument is to provide the absolute certainty Descartes sought, then each of those four premises must be absolutely certain. But we have seen that premise 1, far from being absolutely certain, is highly problematic. Premise 2 involves the rather fuzzy notion of degrees of reality. Premise 7, which says that the cause of an idea's objective or representational reality must (ultimately) be the formal or nonrepresentational reality of something, is somewhat difficult to grasp, though it does seem plausible on reflection, as the analogy of mirror images given in the previous section may have brought out. As for (5), its indubitability was famously challenged by Hume on the grounds that one can conceive of something springing into existence without any cause, that is, of an uncaused beginning of existence. E. M. Curley responds to Hume's challenge this way:

Admittedly I can *conceive* of something springing into existence *ex nihilo*. But I cannot believe that this ever happens.²⁹

This may well be a persuasive defense of (5): Can you seriously believe that something has or will ever come into existence, with no cause or explanation whatsoever?

Perhaps, then, it would not be unreasonable to take (5) and (7) as basic premises; and perhaps (2) can pass muster, as well, despite its unclarity. But one can surely wonder whether those premises are absolutely certain. Furthermore, the problematic status of (1), the precontainment principle, remains a major drawback for this solution.

²⁹ Curley, *Descartes Against the Skeptics*, p. 131.

The other possible solution would be to scrap the subargument and to take (8) as itself a basic premise. Then the entire argument, with its steps renumbered to reflect its shorter length, would reduce to this:

1. The cause of an idea must have as much formal reality as the idea contains objective reality (basic premise).
2. Only a perfect God has as much formal reality as my idea of God contains objective reality (basic premise).
3. The cause of my idea of God is a perfect God (from 1 and 2).
4. A perfect God exists (from 3).
5. To deceive is an imperfection (basic premise).
6. God is not a deceiver (from 4 and 5).
7. If my clear and distinct perceptions could be false, then God would be a deceiver (premise, based on the assent-compellingness of clear and distinct perceptions).
8. My clear and distinct perceptions cannot be false, i.e., whatever I perceive clearly and distinctly must be true (from 6 and 7).

To see the strength of this approach, it is worth quoting once again a passage from Cottingham's *Descartes*, including this time an important observation that Cottingham makes at the very beginning of the passage:

[T]he Causal Adequacy Principle actually seems on *firmer* ground when applied to ideas than when applied to the physical world [emphasis in the original]. Recall Descartes' example of the "highly intricate machine"; and for the sake of simplicity let us follow Descartes' own comparison between ideas and pictures, and consider the case of a drawing or a picture rather than an idea. Suppose a five-year-old child produces a highly complicated design for a computer—a design which we know could only be produced by a highly skilled mathematician with a mental age vastly superior to the child's. The fact that the design is only a drawing and not an actual computer does not block the causal question: the representative or 'objective' intricacy of the design still has to be accounted for. Of course, the child might simply have copied down the drawing from a book. But this simply pushes the argument one stage further back. We are, it seems, justified in asserting that somewhere along the line of causation there must be an actual entity or being that really does possess sufficient complexity to account for the complexities which are to be found in the design.

And what goes for the drawing goes equally for an idea: complex representational content requires a complex cause.³⁰

What conclusion should we draw from our exploration of the grounds for the core argument's first premise? Perhaps the reasonable conclusion is that the premise is plausible, especially if it is taken as a basic premise rather than defended by the subargument Descartes himself gives for it. But one may still wonder whether the premise is as unshakably certain as Descartes took it to be.

4.2 The Problem of the Cartesian Circle

In this section, we turn to one of the most controversial issues in Descartes scholarship: the problem of the Cartesian Circle. This problem is not an objection to Descartes's specific argument for God's existence in the *Third Meditation*; it is an objection to his overall strategy in that *Meditation*. Remember what his purpose was: he wanted to establish or vindicate his principle that whatever I perceive clearly and distinctly is true; that is, to show that clear and distinct perceptions are trustworthy. For on the one hand, he found them to be assent-compelling, but on the other hand, he found that whenever he thought about an omnipotent God, he had to admit that such a God could easily deceive him even about his clearest and most distinct perceptions. To escape from this oscillation, he declared that he must find out whether God exists and can be a deceiver. So he gave an argument, meant to show that God exists and is not a deceiver, and that clear and distinct perceptions are therefore trustworthy. But suppose we ask Descartes: what assures you that the premises of your argument are true and that its steps are valid? The only possible answer for Descartes (and the one he in fact gave) is that he clearly and distinctly perceives the premises to be true and the steps to be valid. But this uses his criterion of truth in the very argument that is supposed to establish that he can safely use it. He is assuming the very thing he set out to prove, namely that his clear and distinct perceptions are true. So the overall argument of *Meditation III* is viciously circular.

In Descartes's own lifetime, this now-famous problem was raised in both the Second and Fourth sets of *Objections* to the *Meditations*. The clearest statement of the problem was the one in the fourth set, written by Antoine Arnauld. Arnauld wrote:

³⁰ Cottingham, *Descartes*, pp. 52–53.

I have one further worry, namely how the author avoids reasoning in a circle when he says that we are sure that what we clearly and distinctly perceive is true only because God exists.

But we can be sure that God exists only because we clearly and distinctly perceive this. Hence, before we can be sure that God exists, we ought to be able to be sure that whatever we perceive clearly and evidently is true. (CSM II 150, SPW 142, M 106, AT VII 214)

Here Arnauld is saying that in order for Descartes to be in a position to give a proof of God's existence, he must already know that whatever he perceives clearly and distinctly is true; for if Descartes does not already know that he can trust his clear and distinct perceptions, then he cannot be sure that the premises of his proof of God's existence are true, or that its steps are logically valid. Hence, Arnauld concludes, Descartes ought not to have held that his criterion of truth requires a divine vindication. He should have held that this criterion is known to be fully reliable even before God's existence is known. Arnauld does not mention another option, which would have been to hold, contrary to what Descartes's text seems to say, that God's existence can be known before knowing that whatever one perceives clearly and distinctly is true.

Recent Descartes scholars have formulated the problem in the following way.³¹ Descartes seems to hold both:

(1) I can know that *whatever I clearly and distinctly perceive is true* only if I first know that *God exists and is no deceiver*

and

(2) I can know that *God exists and is no deceiver* only if I first know that *whatever I clearly and distinctly perceive is true*.

Statement (1) is implied by what Descartes says is the purpose of the *Meditation III* argument for God's existence, namely, to prove that *whatever I clearly and distinctly perceive is true*. To see why Descartes seems also

³¹ This formulation was first suggested by Willis Doney in "The Cartesian Circle," 324–38, and subsequently adopted by James Van Cleve in his masterful "Foundationalism, Epistemic Principles, and the Cartesian Circle" in 1979. Both of these essays are reprinted in Willis Doney ed., *Eternal Truths and the Cartesian Circle: A Collection of Studies* (New York: Garland, 1987).

to be committed to (2), we need only ask: what entitles him to be sure that the premises of his argument for God's existence are true, and to be sure that the inferences drawn from those premises are valid? The only possible answer seems to be this: that he clearly and distinctly perceives those premises to be true and those inferences to be valid. Furthermore, the texts show that this answer is the one Descartes would give. In the *Third Meditation*, he says that the premises of his argument for God's existence—specifically, the Causal Adequacy Principle and the principle that to deceive is an imperfection—are known “by the natural light” (CSM II 28, 35; SPW 91, 98; AT VII 40, 52). But in the *Principles of Philosophy*, he says that

[T]he light of nature or faculty of knowledge which God gave us can never encompass any object which is not true in so far as it is indeed encompassed by this faculty, that is, in so far as it is clearly and distinctly perceived. (CSM I 203, SPW 170, AT VIIIA 16)

This passage implies that the “light of nature” or “natural light” is the very power or faculty of the mind by which clear and distinct perceptions are obtained. Thus, when Descartes says that the premises of his theological argument are known “by the natural light,” he is saying that they are known by being clearly and distinctly perceived, and so, it seems, committing himself to (2). But now, if (1) and (2) above are both true, then Descartes can never know either that *whatever he clearly and distinctly perceives is true* or that *God exists and is no deceiver*; for to know either of these propositions, he would have to know the other one first: but he cannot know them both first! Yet, by the end of the *Third Meditation*, Descartes does claim to know both propositions. The problem is, How is this possible, given the evidence that Descartes is committed to both (1) and (2)?³²

To solve this problem, one would have to show either that Descartes is not really committed to (1), or that he is not really committed to (2). Two different ways of doing this would be to show that appearances to the contrary notwithstanding, Descartes really holds either that:

A. I can know that *whatever I clearly and distinctly perceive is true* before knowing that *God exists and is no deceiver*

³² This analysis is from Van Cleve, “Foundationalism, Epistemic Principles, and the Cartesian Circle,” p. 57.

or that

B. I can know that *God exists and is no deceiver* before knowing that *whatever I clearly and distinctly perceive is true*.

Since strategy A says, in effect, that no divine vindication of the criterion of truth is needed, let us call it the “vindication-not-needed strategy.” Since strategy B says, in effect, that the criterion of truth is not needed to prove God’s existence and veracity, let us call it the “criterion-not-needed strategy.” In the next two subsections, we shall examine recent defenses of each strategy.

4.2.1 *The Restriction of the Doubt to Past Clear and Distinct Perceptions Defense*

This defense, which for brevity’s sake we shall henceforth refer to as “the restriction defense,” claims that Descartes doesn’t really accept thesis (1), that I can know that *whatever I clearly and distinctly perceive is true* only if I first know that *God exists and is no deceiver*. For the proof of God’s existence and veracity is not required to guarantee that present clear and distinct perceptions are true, but only in cases where I am relying on *past, remembered clear and distinct perceptions*. In other words, what we may call “the divine guarantee” is not needed to guarantee the proposition that *whatever I clearly and distinctly perceive is true*, but only to guarantee this proposition: “Whatever I remember having clearly and distinctly perceived is true.” The restriction defense should not be confused with the now-discredited, so-called memory defense, which turns on the idea that memory is fallible but that God guarantees that memories of (only) clear and distinct perceptions are infallible; for it does not turn on any doubt about the reliability of memory, but only on the *pastness* of the perceptions, which are assumed to be correctly remembered.³³ The basic assumption of the restriction defense is, as Gary Hatfield puts it, that

³³ The main proponent of the memory thesis was Willis Doney in “The Cartesian Circle.” Later, however, Doney advanced a different solution to the problem of the circle, in his “Descartes’s Conception of Perfect Knowledge,” 387–403 (also reprinted in his (ed.) *Eternal Truths and the Cartesian Circle*). The most powerful critique of the memory defense is Harry G. Frankfurt, “Memory and the Cartesian Circle,” 504–11. The memory defense and the reasons for its failure were discussed in the first edition of this book, but will not be covered in this edition.

"the reliability of clear and distinct perception was never itself really placed in doubt, merely our ability to remain convinced of its reliability when we aren't having such perceptions."³⁴

To see how the divine guarantee is supposed to work according to the restriction defense, let us use Descartes's own illustration in *Meditation V*. Suppose I have worked through a proof that the three angles of a triangle equal two right angles. Schematically, the proof looks as shown in Figure 3.6.

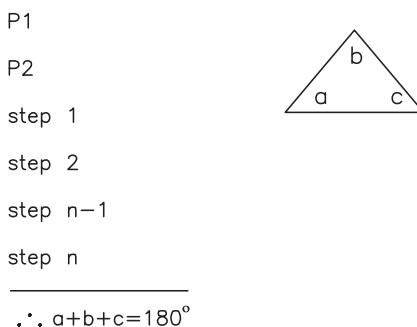


Figure 3.6

So long as I hold the entire proof before my mind and clearly and distinctly perceive each premise and step, I do not need the divine guarantee, because I am not relying on *past* clear and distinct perceptions. The two cases where I need it are (1) when I remember that I proved that $a + b + c = 180^\circ$ but I am no longer attending to the proof at all and so not currently clearly and distinctly perceiving any part of it, and (2) when the proof is too long or complicated for me to grasp all of it at once, so that by the time I get to (say) step n, I am no longer clearly and distinctly perceiving any of the earlier premises or steps. Since the restriction defense denies (1)—that is, says that I can know that *whatever I clearly and distinctly perceive is true* before (without) knowing that *God exists and is no deceiver*—it conforms to the vindication-not-needed strategy.

There are several passages in Descartes's works that suggest that he means to restrict the doubt to only past clear and distinct perceptions. In the passage from *Meditation V* where he offers the above illustration, he says,

³⁴ Hatfield, *Descartes and the Meditations*, pp. 170–71.

Admittedly my nature is such that so long as I perceive something very clearly and distinctly I cannot but believe it to be true. But my nature is also such that I cannot fix my mental vision continually on the same thing, so as to keep perceiving it clearly; and often the memory of a previously made judgment may come back, when I am no longer attending to the arguments which led me to make it. And so other arguments can now occur to me which might easily undermine my opinion, if I were unaware of God; and I should thus never have true and certain knowledge about anything, but only shifting and changeable opinions. For example, when I consider the nature of a triangle, it appears most evident to me, steeped as I am in the principles of geometry, that its three angles are equal to two right angles; and so long as I attend to the proof, I cannot but believe this to be true. But as soon as I turn my mind's eye away from the proof, then in spite of still remembering that I perceived it very clearly, I can easily fall into doubt about its truth, if I am unaware of God. (CSM II 48, SPW 109, AT VII 69–70)

Now, however, I have perceived that God exists, and at the same time I have understood that everything else depends on him, and that he is no deceiver; and I have drawn the conclusion that everything which I clearly and distinctly perceive is of necessity true. Accordingly, even if I am no longer attending to the arguments which led me to judge that this is true, as long as I remember that I clearly and distinctly perceived it, there are no counter-arguments which can be adduced to make me doubt it, but on the contrary I have true and certain knowledge of it. And I have knowledge not just of this matter, but of all matters which I remember ever having demonstrated, in geometry and so on. (CSM II 48, SPW 109–110, AT VII 70)

Furthermore, when Descartes tries to refute the charge that his reasoning is circular, he seems again to resort to the restriction defense. Thus, the author(s) of the Second set of *Objections* had said:

You are not yet [at the start of *Meditation III*] certain of the existence of God, and you say that you are not certain of anything, and cannot know anything clearly and distinctly until you have

achieved clear and certain knowledge of the existence of God. It follows from this that you do not yet clearly and distinctly know that you are a thinking thing, since, on your own admission, that knowledge depends on the clear knowledge of an existing God; and this you have not yet proved in the passage where you draw the conclusion that you clearly know what you are. (CSM II 89, SPW 139, M 102, AT VII 124–5)

Descartes's reply to this was:

When I said that we can know nothing for certain until we are aware that God exists, I expressly declared that I was speaking only of knowledge of those conclusions which can be recalled when we are no longer attending to the arguments by means of which we deduced them. (CSM II 100, SPW 139, M 103, AT VII 140)

Later, in replying to Arnauld's previously-quoted version of the circularity charge, Descartes refers back to this response, saying:

Lastly, as to the fact that I was not guilty of circularity when I said that the only reason we have for being sure that what we clearly and distinctly perceive is true is the fact that God exists, but that we are sure that God exists only because we perceive this clearly: I have already given an adequate explanation of this point in my reply to the Second Objections . . . where I made a distinction between what we in fact perceive clearly and what we remember having perceived clearly on a previous occasion. To begin with, we are sure that God exists because we attend to the arguments which prove this; but subsequently it is enough for us to remember that we perceived something clearly in order for us to be certain that it is true. This would not be sufficient if we did not know that God exists and is not a deceiver. (CSM II 171, SPW 142–3, M 106, AT VII 245–6)

Although the above passages suggest that Descartes adopted the restriction defense, that defense faces serious objections, both philosophical and textual. The main philosophical objection is that it seems to imply that every time one wishes to use the divine guarantee, one must

go through the proof of God's existence and hold it before one's mind. To see this, notice first that on the restriction defense, the proof of God's existence must, of course, itself not rely on past clear and distinct perceptions. Rather, it must be possible to grasp the entire proof at once, instead of having to rely on premises or steps that were clearly and distinctly perceived only in the past; for if such premises or steps were required, then one could never be certain of the conclusion, since all past clear and distinct perceptions are suspect until that conclusion has been established. Now Descartes did, in fact, hold that his proof of God's existence can, after some practice, be grasped all at once, in a single mental "intuition." The difficulty, however, concerns proofs that are too long to be grasped all at once, such as lengthy mathematical proofs. It is precisely in those cases, according to the restriction defense, that the divine guarantee is needed. But how would one then use the guarantee? Would it be enough to appeal to the fact that one previously proved God's existence from clearly and distinctly perceived premises and steps? No; for then one would be defending the reliability of some past clear and distinct perceptions (e.g., those of the first several steps of a mathematical proof) by means of an appeal to other past clear and distinct perceptions (i.e., those of all the steps of the theological argument)—which is circular if the reliability of past clear and distinct perceptions is in doubt. To use an example of Ludwig Wittgenstein's (1889–1951), it would be like trying to establish the truth of a newspaper story by consulting another copy of the same newspaper. So to use the divine guarantee, one must go through the proof of God's existence all over again. One must do this each time one uses the guarantee, which is, at least, an awkward result. Furthermore, Descartes himself denies that such repeated rehearsals of the proof are needed: in the second passage quoted above from *Meditation V*, he says,

[So] long as I remember that I clearly perceived it [i.e., that God exists and is no deceiver] . . . I have true and certain knowledge of it. And I have knowledge not just of this matter, but of all matters which I remember ever having demonstrated, in geometry and so on. (CSM II 48, SPW 109–110, AT VII 70)

Here Descartes explicitly says that, to use the divine guarantee, one need only remember having clearly and distinctly perceived (a proof of) God's existence. As we have just seen, this would be manifestly circular if the

function of the divine guarantee were to insure the reliability past clear and distinct perceptions.

The other difficulty for the restriction defense is a textual one: it simply does not square with what Descartes says in the *Meditations*. Recall the important passage, near the beginning of the *Third Meditation*, where Descartes was oscillating between certainty about present clear and distinct perceptions, and doubt about the reliability of his cognitive faculties. There Descartes said, referring to things that are “very simple and straightforward in arithmetic or geometry, for example that two and three added together make five, and so on”:

[T]he only reason for my later judgment that they were open to doubt was that it occurred to me that perhaps some God could have given me a nature such that I was deceived even in matters which seemed most evident. And whenever my preconceived belief in the supreme power of God comes to mind, I cannot but admit that it would be easy for him, if he so desired, to bring it about that I go wrong even in those matters which I think I see utterly clearly with my mind's eye. (CSM II 36, SPW 87–88, AT VII 36)

Here Descartes does not even mention past clear and distinct perceptions. Rather, his point is that an omnipotent God could easily make him go wrong even about the things that he *currently* perceives with the utmost clarity and distinctness and cannot doubt while he is attending to them. So, Descartes here carries his doubt much further than the restriction defense allows. Of course it is true that at any particular time, only past clear and distinct perceptions can be doubted, since present clear and distinct perceptions are assent-compelling. But this does not solve the problem; for it does not mean that at any time after having clearly and distinctly perceived that *God exists and is not a deceiver*, I cannot doubt that proposition. So the fact that I earlier clearly and distinctly perceived it does not allow me to escape from the oscillation between being certain of this proposition and doubting it (and all the ones that depend on it).

The basic problem for the restriction defense can be put like this. To say that past clear and distinct perceptions can be doubted is to say that even though I clearly and distinctly perceived some proposition in the past, that proposition may have been false. This is to grant that the

fact that a proposition is clearly and distinctly perceived does not entail that the proposition is true; it is to open a logical gap between clarity and distinctness and truth. But once that gap is opened, how can it be closed? The restriction defense's only answer is to point to the fact that we are incapable of doubting a present clear and distinct perception—to the assent-compellingness of present clear and distinct perceptions. But this assent-compellingness is merely a psychological fact about us; it does not in itself go any way toward showing that clear and distinct perceptions must be true. Therefore, it cannot possibly close the gap: once it is granted that a proposition that was clearly and distinctly perceived in the past might be false, there is no way to prevent doubt about that proposition from recurring whenever it is not being clearly and distinctly perceived, regardless of how strongly we were convinced of its truth at the moment that it was being clearly and distinctly perceived.

4.2.2 *The General Rule Defense*

According to the criterion-not-needed-strategy, Descartes holds that his criterion of truth does require vindication, and the purpose of his theological argument is to supply that vindication. However, the criterion of truth is not necessary for the theological argument itself. Note, then, that the point of the label “criterion-not-needed” is not that Descartes doesn’t need his criterion of truth to complete his overall program but only that the criterion is not needed *for his proof of God’s existence and veracity*.

The obvious question for the criterion-not-needed-strategy is, How can Descartes legitimately appeal to the clarity and distinctness of the premises and steps of his proof of God’s existence and veracity if he admits that his criterion of truth does need a vindication—one that is not available until after the proof is complete and a nondeceiving God is known to exist? In other words, how can Descartes both appeal to the clarity and distinctness of his proof of God’s existence’s premises and steps and yet hold that the clarity and distinctness criterion is not needed for that proof?

Let us examine an answer first proposed by Anthony Kenny and subsequently adopted, with minor modifications, by James Van Cleve and by Bernard Williams. This answer is that the divine guarantee that the proof of God’s existence provides is needed only for knowledge of the *general*

rule that whatever I perceive clearly and distinctly is true; it is not needed for the individual, particular clear and distinct perceptions appealed to in the proof (or for any other individual or particular clear and distinct perceptions). Let us call this answer “the general rule defense.”³⁵ Our thesis in this subsection will be that although the general rule defense is ingenious, it is ultimately unsuccessful.

Notice that both the restriction defense and the general rule defense hold (albeit for very different reasons) that the divine guarantee is not required for presently occurring clear and distinct perceptions—the restriction defense because the divine guarantee is needed only for past clear and distinct perceptions, the general rule defense because it is needed only for the *generalization* that whatever I perceive clearly and distinctly is true. Thus, on both defenses, the clear and distinct perceptions that I have when I go through the proof of God’s existence and veracity are immune from all doubt: there is no circle.

What textual support is there for the general rule defense? Mainly that Descartes never says that he can doubt particular clear and distinct perceptions, such as his perceptions that $2 + 3 = 5$, *cogito ergo sum*, or *not-(p and not-p)*. On the contrary, he steadfastly maintains that such perceptions are “assent-compelling”—that is, he *cannot* doubt them *while they actually occur*. Thus, near the beginning of *Meditation III*, immediately after admitting that an omnipotent God could easily make him err “even in those matters which I think I see utterly clearly with my mind’s eye,” he insists that he nevertheless cannot doubt such matters while actually attending to them:

Yet when I turn to the things themselves which I think I perceive clearly, I am so convinced by them that I spontaneously declare: let whoever can do so deceive me, he will never bring it about that I am nothing, so long as I continue to think that I am something; or make it true at some future time that I have never existed, since it is now true that I exist; or even bring it about that two and three added together are more or less than five, or

³⁵ The general rule defense was first suggested by Kenny, *Descartes*, chapter 8, and more perspicuously presented in part I of his “The Cartesian Circle and the Eternal Truths,” 685–700. Van Cleve endorses it (with some modifications and elaborations) in his “Foundationalism, Epistemic Principles, and the Cartesian Circle,” pp. 66–69. See Van Cleve’s n. 28 for one indispensable clarification of Kenny’s position. Williams takes a similar position in his *Descartes: The Project of Pure Enquiry*, p. 203.

anything of this kind in which I see a manifest contradiction.
(CSM II 25, SPW 88, AT VII 36)

In a passage in *Meditation V* where Descartes is referring to mathematical truths whose proofs he clearly and distinctly perceives, he makes the same point this way:

I have already amply demonstrated that everything of which I am clearly aware is true. And even if I had not demonstrated this, the nature of my mind is such that I cannot but assent to these things, at least so long as I clearly perceive them. (CSM II 45, SPW 106, AT VII 65)

Notice that Descartes here explicitly says that his clear and distinct perceptions would be assent-compelling, even if he had *not* established his general rule that clear and distinct perceptions must all be true. A couple of pages later, he puts the point very concisely:

[M]y nature is such that so long as I perceive something very clearly and distinctly I cannot but believe it to be true. (CSM II 48, SPW 109, AT VII 69)

Then he gives an example:

For example, when I consider the nature of a triangle, steeped as I am in the principles of geometry, it appears most evident to me that its three angles are equal to two right angles; and so long as I attend to the proof, I cannot but believe this to be true. (CSM II 48, SPW 109, AT VII 69)

Finally, there is an exceptionally strong statement of the point in Descartes's *Reply to the Second Set of Objections*. Referring to "the clear perceptions of the intellect," Descartes says:

Now some of these perceptions are so transparently clear and at the same time so simple that we cannot ever think of them without believing them to be true. The fact that I exist so long as I am thinking, or that what is done cannot be undone, are examples of truths in respect of which we manifestly possess this kind of

certainty. For we cannot doubt them unless we think of them; but we cannot think of them without at the same time believing that they are true . . . Hence we cannot doubt them without at the same time believing that they are true; that is, we can never doubt them. (CSM II 104, M 105, SPW 141–2, AT VII 145–6)

Here Descartes claims that some truths, such as the *cogito* and the law of noncontradiction, are so simple in their content that one cannot even think of them without clearly and distinctly perceiving them and therefore being compelled to assent to them. So, since one cannot doubt a proposition without thinking of it, these truths can never be doubted at all.

What Descartes does admit, however, is that when he thinks about an omnipotent God, he has to concede that such a God could make him go wrong even about those things that seem most evident to him. Now the key claim of the general rule defense is that this admission does *not* contradict Descartes's claim that particular clear and distinct perceptions cannot be doubted while they occur; for according to the general rule defense, Descartes's admission has a specific meaning. It means that, when Descartes thinks of God's omnipotence, he can doubt the *generalization* that whatever he most clearly and distinctly perceives is true. From this, it does not follow that he can doubt any particular clear and distinct perception that he is having. From

I can doubt that for every proposition *p*, if I clearly and distinctly perceive that *p*, then it is true that *p*

it does not follow that

There is a proposition *p* such that I clearly and distinctly perceive that *p*, and I can doubt that *p*.

As Van Cleve puts it, “[Descartes] might be uncertain of the general connection between clear and distinct perception and truth, yet certain of every proposition [he] . . . clearly and distinctly perceive[s].”³⁶

Anthony Kenny puts the matter this way. Before proving God's existence and veracity, Descartes is prepared to admit

³⁶ Van Cleve, “Foundationalism, Epistemic Principles, and the Cartesian Circle,” p. 67. See also p. 100, above.

- (1) For some proposition p , I clearly and distinctly perceive that p , but maybe $\text{not-}p$.

However, Descartes is not prepared to admit any particular case of (1). For example, he is not prepared to admit

- (a) I clearly and distinctly perceive that $2 + 3 \text{ is } 5$, but maybe $2 + 3 \text{ is not } 5$

or

- (b) I clearly and distinctly perceive that *if I am thinking then I exist*, but maybe *I am thinking yet I do not exist*.³⁷

As Kenny points out, it is perfectly possible for a person to accept (1) without accepting any specific case of (1). This is much like admitting that some of one's beliefs are false without admitting (paradoxically) that any particular one of them is false. Almost all people, unless they are extremely arrogant, would probably admit

- (2) For some p , I believe p ; but $\text{not-}p$.

But hardly anyone would admit any particular case of (2), such as "I believe that it is raining, but it is not raining"; for as the British philosopher G. E. Moore (1873–1958) pointed out, this sentence—or any other sentence of the form "I believe that p , but $\text{not-}p$ "—is highly paradoxical. (Such sentences are all cases of what philosophers call "Moore's Paradox.") Likewise, then, Descartes can admit (1) without accepting paradoxical statements like (a) and (b).

After proving God's existence and veracity, however, Descartes no longer admits (1). For now he has established the opposite of (1), namely the general rule that

- (3) For all p , if I clearly and distinctly perceive that p , then p is true.

³⁷ The material in this and the next two paragraphs closely follows Kenny, "The Cartesian Circle and the Eternal Truths," pp. 687–90.

Kenny also formulates the matter this way. Even before proving God's existence and veracity, the assent-compellingness of clear and distinct perceptions means that the following statement is true of me:

- (4) For all p , if I clearly and distinctly perceive that p , then I cannot doubt that p .

On the other hand, because of the general doubt about the reliability of clear and distinct perception based on the possibility of a deceiving God, it is false (before proving God's existence and veracity) to say

- (5) I cannot doubt that: for all p , if I clearly and distinctly perceive that p , then p is true.

But after the proof of a perfect God and the consequent removal of this generalized doubt, (5) also becomes true of me.

Despite what Kenny says, however, you might think that Descartes's admission that before knowing God's existence and veracity he can doubt even those things that he most clearly and distinctly perceives must contradict at least his assertion in the second set of *Replies* that some very simple truths, like the *cogito* and the law of noncontradiction, can never be doubted at all. On one interpretation of Descartes's admission, it would seem to contradict that assertion. If the admission is interpreted to mean that I can doubt whether *any* of the propositions that I clearly and distinctly perceive are true, then, given also that the *cogito* and the law of noncontradiction are propositions that I clearly and distinctly perceive, it seems to follow that I can doubt them—which contradicts the assertion that I can never doubt them. But there is another way to interpret Descartes's admission, on which it does not lead to contradiction. The admission can be interpreted to mean that I can doubt whether *all* the propositions that I clearly and distinctly perceive are true, that is, that I can admit that some or most of them might be false. This certainly does not mean that I can doubt the truth of *every* proposition that I clearly and distinctly perceive: there may be some exceptional ones that I can never doubt. So the admission does not contradict the claim that I cannot doubt the *cogito* or the law of noncontradiction. Now according to the general rule defense, what can be doubted before proving God's existence and veracity is just this general rule that all clearly and distinctly perceived propositions are true; and this is quite compatible with saying that some

specific clearly and distinctly perceived propositions can never be doubted. Thus, it appears that Descartes's admission, if interpreted in accordance with the general rule defense, is consistent even with his assertion that some propositions can never be doubted.

Now that we have presented the general rule defense, it is time to examine it critically. We shall raise two different difficulties for this defense, and argue that while it can deal with the first difficulty, it succumbs to the second one.

The first difficulty is that if knowledge of the general rule that whatever I perceive clearly and distinctly is true is not needed for the proof of God's existence and veracity, then presumably it is not needed *for any other arguments* whose premises and steps are clear and distinct, such as mathematical proofs. So what is the use of the general rule? Can Descartes avoid circularity only at the price of rendering the rule (and its divine vindication) useless?

An illuminating answer to this question has been offered by Bernard Williams in his *Descartes*. Essentially, Williams's answer is that without the general rule, one would be limited to momentary episodes of certainty: one would be certain of only those propositions that one was currently perceiving clearly and distinctly and so currently unable to doubt. But as soon as one stopped attending to those propositions, one could wonder whether what was previously perceived clearly and distinctly is true. But this means that one would never really *know* or, Williams suggests, even *believe* anything; for both knowledge and belief are on-going states rather than momentary episodes. As Williams puts it, Descartes

finds that he has encountered a number of propositions which are irresistible [i.e., believed if even thought of, like the *cogito* and the law of noncontradiction]. These encounters, however, have not given [Descartes] any knowledge. . . . There is even a sense, and an important one, in which they have not given him any beliefs. The perceptions of these propositions which have occurred so far, and in which they revealed their irresistibility, do not satisfy what is virtually a formal requirement on knowledge or (in a full sense) belief, that it should be an on-going dispositional state. . . . The clear and distinct perceptions . . . which [Descartes] has had are all time-bound, in the sense that he may at one time clearly and distinctly perceive that *p* is true, but not at another.³⁸

³⁸ Williams, *Descartes: The Project of Pure Enquiry*, pp. 200–201.

Williams goes on to say that, faced with the prospect of having only such momentary episodes of certainty, one has three choices: (1) just give up the search for stable, certain beliefs, (2) freeze one's attention permanently on just one clear and distinct perception, or (3) adopt some "acceptance-rule" that would, as it were, promote momentary episodes of certainty into full-fledged beliefs. He then interprets Descartes as taking the third option, as "admitt[ing] some acceptance-rule for beliefs which are on-going and not time-bound as the clear and distinct perceptions are." The rule that Descartes adopts, he says, is "Accept as on-going beliefs just those propositions which are at any time clearly and distinctly perceived to be true."³⁹

In his *Descartes*, John Cottingham adopts a view very similar to Williams's. Before quoting a passage from Cottingham's book, however, we should note that he interprets Descartes's claim that present clear and distinct perceptions cannot be doubted differently than we have done. He takes it to mean not merely that present clear and distinct perceptions are assent-compelling (i.e., cannot be doubted while they occur), but also that they are "self-guaranteeing" (i.e., *known to be true just by being had*). There is, indeed, a deep issue at stake in this difference of interpretation, to which we shall come shortly. But the point to note for now is that having made the claim that clear and distinct perceptions are self-guaranteeing, Cottingham faces, all the more urgently, the question, What is the use of the divine guarantee? As he puts it, "How do we construe Descartes' frequent assertions that God is the source and guarantor of all knowledge?" The answer, he says,

lies in the very temporary nature of the self-guaranteeing flashes of intuition which the meditator enjoys. The guaranteed recognition of truth lasts, for a given proposition, only so long as the meditator holds that proposition in front of his mind; as soon as his attention wanders, even for a moment, the guarantee vanishes . . . Once God's existence is established, however, then we have the possibility of progressing beyond such isolated flashes of cognition and building up a systematic body of knowledge. As Descartes puts it, we can move from *cognitio* (mere cognition) to *scientia* (stable knowledge).⁴⁰

³⁹ Williams, *Descartes: The Project of Pure Enquiry*, pp. 201–02.

⁴⁰ Cottingham, *Descartes*, p. 70.

This is essentially the same way of accounting for the use of the general rule that Williams proposes.

Let us turn to the second difficulty for the general rule defense. This difficulty concerns the most crucial issue for that defense: Does it really provide a way of showing that I can know that God exists and is no deceiver before knowing that whatever I clearly and distinctly perceive is true, thereby solving the problem of the circle? Despite the ingenuity of the general rule defense, we may question whether it really does solve the problem. The proponents of this defense would presumably say (as, in fact, Kenny does say) that *before* God's existence and veracity are proved, the general principle that whatever I perceive clearly and distinctly is true can be doubted. Now what is it to doubt this principle? It is not to doubt an isolated and perhaps insignificant proposition. Rather, doubting the general principle must consist in thinking that even when I am having a clear and distinct perception, which admittedly I cannot doubt while I am having it, I may nevertheless be mistaken: the proposition that I clearly and distinctly perceive may actually be false. Or, to put it in a way that uses verb tenses so as to highlight the fact that I cannot doubt a proposition *at the time I am clearly and distinctly perceiving it*, doubting the general principle must consist in thinking, "Even when I *was* having a clear and distinct perception, which admittedly I could not doubt at the time I was having it, I may nevertheless have been mistaken: the proposition that I *was* then clearly and distinctly perceiving may actually have been false." There seems to be no other way to understand what doubting the principle that whatever I perceive clearly and distinctly is true amounts to. Notice, then, that the general rule defense is *incompatible* with the view that clear and distinct perceptions are not only assent-compelling, but also self-guaranteeing. For the doubt of the general rule to have any content, it must at least allow for doubt of past clear and distinct perceptions and so for the possibility that a proposition that was clearly and distinctly perceived was nevertheless false and, therefore, not known to be true. To put it differently, the general rule defense must allow a doubt concerning the *faculty* for clear and distinct perception—for a doubt concerning the very reliability of human reason.

Now according to the general rule defense, Descartes purports to remove this doubt by giving a proof of the general rule itself, turning on

God's existence and veracity. But if, prior to knowing God's existence and veracity, I must admit that I could be mistaken even about my clearest and most distinct perceptions, then how can such a proof provide any lasting assurance? Admittedly, I will be unable to doubt the proof so long as I attend to it. If I can grasp the entire proof at once, including the step leading to the final conclusion that whatever I perceive clearly and distinctly is true, then there will be a time or times when I cannot doubt that general principle. But the moment I turn away from the proof, the doubt can recur. I can then say to myself:

I remember constructing an argument showing that whatever I perceive clearly and distinctly is true, using certain premises and making certain logical steps that I perceived most clearly and distinctly and could not doubt at the time. One of those premises was even the premise that God is not a deceiver, from which I inferred that he would not allow my clear and distinct perceptions to be false. But when I stop attending to this argument and think instead of the supreme power of God, I must admit that he could easily cause me to go wrong even about those things that I perceive utterly clearly, including of course the premises and the steps of the very argument which was supposed to show that this is not so. So, I cannot be certain even now that there is a God who would not deceive me, or that whatever I clearly and distinctly perceive is true.

Thus, it seems that Descartes cannot emerge from the oscillation between his certainty about presently occurring clear and distinct perceptions, and his general doubt about the reliability of even his clearest and most distinct perceptions, from which his theological argument was supposed to free him.

Consider an analogy. Suppose there were a drug that works as follows. While its effect lasts, various propositions are indubitable for the drug-taker: he cannot possibly doubt them. As long as the drug's effect lasts, these propositions are assent-compelling. Furthermore, one of the propositions which is thus indubitable while the drug is working is the proposition that whatever I perceive while under the drug's influence is true. So while I am under the drug's influence, I cannot doubt that perceiving a proposition while under the drug's influence guarantees truth: I am certain that it does. Now, how could this temporary certainty permanently

remove any antecedent doubts I might have had about whether being under the drug's influence guarantees truth? And why should it prevent me from doubting this after the drug's effect has worn off? You may want to complain that it's unfair to compare clear and distinct perception to a drug; for the drug could well be a cause of error, illusion or delusion, whereas the clear and distinct perceptions of the intellect are our best or most reliable way of accessing truth. But this complaint would miss the point of the analogy. This is that if we assume that the most that can be said in favor of clear and distinct perceptions (before proving God's existence) is that they are assent-compelling, then we are by that very assumption granting that their deliverances might be no better than the effects of a drug. So these deliverances, no matter how assent-compelling they may be when clear and distinct perceptions are occurring, cannot provide any guarantee that clear and distinct perceptions are true, even if one of the deliverances of clear and distinct perception is that clear and distinct perceptions themselves are true.

Of course, if Descartes is right to hold, in the Second set of *Replies*, that some propositions are so simple that they can never be thought of without being grasped and so obvious that they cannot be grasped without being believed, then there will be a few, exceptional propositions (e.g., the *cogito* and the law of noncontradiction) that are completely indubitable. But these propositions do not include the propositions "God exists and is no deceiver" and "Whatever I clearly and distinctly perceive is true." So Descartes will be unable to emerge from his vacillation about the truth of all things beyond the few exceptional propositions.

Indeed one philosopher, Harry Frankfurt, has in effect suggested that without the divine guarantee of clear and distinct perceptions, Descartes would not claim to know that even those few exceptional propositions are true. Frankfurt maintains that the most basic problem of the *Meditations* is the relationship between indubitability and truth.⁴¹ So, he insists that when Descartes claims that a proposition is indubitable, this must never be confused with his claiming that the proposition is true. On Frankfurt's reading, then, we are to see Descartes as countenancing the thought "There are some propositions which I can *never* doubt, but which

⁴¹ See Harry G. Frankfurt, "Descartes' Validation of Reason," 149–56 (reprinted in Doney, ed., *Descartes: A Collection of Critical Essays* (Garden City, Doubleday/Anchor Books, 1967), idem, *Demons, Dreamers, and Madmen* (New York: Bobbs-Merill, 1970), pp. 162–66.

may nevertheless be false." As we shall see in the next subsection, there is indeed a case for saying that Descartes would countenance such a thought. But one may question whether the thought is a coherent one, for does it not amount to saying that there are some propositions that I both can and cannot doubt? Be that as it may, the problem that we have described also concerns the relationship between indubitability and truth; for the problem is, basically, that although some propositions cannot be doubted at certain times (while one is clearly and distinctly perceiving them or their derivation from other currently clearly and distinctly perceived propositions), they can be doubted at other times, and so admitted, at least at those times, to be possibly false. This problem remains even if one holds, against Frankfurt, that Descartes's claim that a few exceptional propositions can never be doubted conflicts with saying that those propositions might be false.

4.2.3 The Radical Doubt of Reason and the Creation of the Eternal Truths

Our critique of the general rule defense has brought us to the heart of the problem of the Cartesian Circle. The goal that Descartes sets for himself at the beginning of the *Third Meditation* is, at bottom, to overcome a doubt about the reliability of human reason (whether it be called "the natural light," the "power of clear and distinct perception," or whatever) by means of an argument or chain of reasoning. To borrow Harry Frankfurt's phrase, Descartes's goal is to provide a "validation of reason."⁴² Now, to overcome a doubt about the reliability of reason by means of a chain of reasoning, one must obviously use reason. But if the reliability of reason is itself doubtful, then how can the results of this (or any other) use of reason be trusted? Isn't Descartes's project of overcoming a doubt about reason by using reason doomed from the start? As David Hume put it:

There is a species of skepticism, antecedent to all study and philosophy, which is much inculcated by Des Cartes and others, as a sovereign preservative against error and precipitate judgement. It recommends an universal doubt, not only of all our former opinions and principles, but also of our very faculties; of whose

⁴² Frankfurt, *Dreamers, Demons, and Madmen*, chap. 15; idem, "Descartes' Validation of Reason."

veracity, say they, we must assure ourselves, by a chain of reasoning, deduced from some original principle, which cannot possibly be fallacious or deceitful. But neither is there any such original principle, which has a prerogative above others, that are self-evident and convincing: or if there were, could we advance a step beyond it, but by the use of those very faculties, of which we are supposed to be already diffident. The Cartesian doubt, therefore, were it ever possible to be attained by any human creature (as it plainly is not) would be entirely incurable; and no reasoning could ever bring us to a state of assurance and conviction upon any subject.⁴³

The twentieth-century Oxford philosopher, H. A. Prichard, put the same point this way:

Descartes' idea of how the doubt has to be allayed, if at all, is obviously mistaken; and we can see that it is mistaken without even considering the actual way in which Descartes considers that he allays it, viz. by discovering what he considers to be a proof that there exists a deity who is no deceiver. For as Descartes is representing the matter, such a proof could only consist in perceiving clearly and distinctly that certain things which he perceives clearly involve a necessity that such a God exists; and as precisely what he is doubtful of is whether even when he perceives something clearly and distinctly he is not mistaken, he will become doubtful, when he reflects on his having this proof, whether here too, he is not mistaken. . . . Descartes, therefore, it is obvious, is setting himself an impossible task; and we can discover it to be impossible without considering his attempt to execute it.⁴⁴

In light of this powerful objection, it may be tempting to regard Descartes's assertions that an omnipotent God could deceive him about even his clear and distinct perceptions as aberrations, that is, as temporary, ill-considered departures from his own main line of thought. It may be tempting to say, with Cottingham, that for Descartes clear and distinct

⁴³ Hume, *An Enquiry Concerning Human Understanding*, Section 12, part I, para. 3.

⁴⁴ H. A. Prichard, "Descartes's Meditations," pp. 146–47.

perceptions are self-guaranteeing (known to be true just by being had). The Cartesian doubt would then be restricted, as it was in the first and second *Meditations*, to sense-perception, leaving the clear and distinct perceptions of reason as “safe” resources upon which to rebuild one’s knowledge.

Although it might be reasonable to contend that this is the course Descartes *should* have taken, it seems factually incorrect to say that it is the course he actually *did* take. There are two different reasons for this.

First, as we have already seen, there are a number of passages in the *Meditations* and in Descartes’s *Replies* to the Second Set of *Objections* where he says that God might have given him a nature such that he goes wrong even about matters that seem most evident to him. Nor are the passages already discussed the only ones of this kind. In his *Replies* to the Second Set of *Objections*, he also says that “an atheist . . . cannot be certain that he is not being deceived on matters which seem to him to be very evident” (CSM II 101, SPW 140, AT VII 141). In the *Principles of Philosophy*, he says that prior to knowing the existence of God, the mind “is still ignorant as to whether it may have been created with the kind of nature that makes it go wrong even in matters which appear most evident” (CSM I 197, SPW 164, AT VIIIA 9–10), and that the proof of God’s existence and veracity “disposes of the most serious doubt which arose from our ignorance about whether our nature might not be such as to make us go wrong even in matters which seemed to us utterly evident” (CSM I 203, SPW 170, AT VIIIA 16). In short, there is compelling textual evidence that Descartes was prepared to extend his doubt even to reason itself.

The second reason for denying that Descartes held clear and distinct perceptions to be self-guaranteeing stems from a doctrine of his that up to now, we have not mentioned. This is his view that God created the principles of logic and mathematics. According to this doctrine, which commentators call “the creation of the eternal truths,” the principles of mathematics and logic depend on God, and he could have made those principles different from what they are. Although Descartes never mentions this extraordinary doctrine in the *Meditations*, he asserts it several times in his correspondence. Here are some representative passages:

The mathematical truths which you call eternal have been laid down by God and depend on Him entirely no less than the rest of His creatures. Indeed to say that these truths are independent of God is to talk of Him as if he were Jupiter or Saturn and to

subject Him to the Styx and the Fates. Please do not hesitate to assert and proclaim everywhere that it is God who has laid down these laws in nature just as a kind lays down laws in his kingdom. (To Mersenne, April 15, 1630, CSMK 23, AT I 145)

As for the eternal truths . . . we must not say that if God did not exist nevertheless these truths would be true; for the existence of God is the first and most eternal of all possible truths and the one form which alone all others proceed. . . . [People] should . . . take the . . . view, that since God is a cause whose power surpasses the bounds of human understanding . . . these truths are therefore something less than, and subject to, the incomprehensible power of God. (To Mersenne, May 6, 1630, CSMK 24–25, AT I 149–50)

You ask me by what kind of causality God established the eternal truths. I reply: by the same kind of causality as he created all things, that is to say, as their efficient and total cause. . . . [God] was free to make it not true that all the radii of the circle are equal—just as free as he was not to create the world. (To Mersenne, May 27, 1630, CSMK 25, AT I 152)

[E]ven those truths which are called eternal—as that the whole is greater than its parts—would not be truths if God had not so established, as I think I wrote you once before. (To Mersenne, May 27, 1638, CSMK 103, AT II 134)

God cannot have been determined to make it true that contradictions cannot be true together, and therefore . . . he could have done the opposite. (To Mesland May 2, 1644, CSMK 235, AT IV 118)

Descartes's doctrine of the creation of the eternal truths is powerful evidence against the view that he held clear and distinct perception to be self-guaranteeing; for this doctrine implies a possible scenario on which, according to Descartes, the very principles that we take to be the most clear and obvious could be false. According to this scenario, (a) God could have made it *false* that $2 + 2 = 4$, or that *not-(p and not-p)*, or that *a whole is greater than any of its parts*, or the like, and yet (b) have made our minds such that we cannot see the sum of two and two as being anything other than four, or the truth of *p* as being anything other than incompatible

with the truth of *not-p*, or a whole as being less than, or equal to, any of its parts, and so forth. As Descartes himself says to Arnauld:

[S]ince everything in truth and goodness depends on [God's] omnipotence, I would not dare to say that God cannot make . . . [it] that one and two should not be three. I merely say that He has given me such a mind that I cannot conceive . . . an aggregate of one and two which is not three, and that such things involve a contradiction in my conception. (To Arnauld, July 29, 1648, CSMK 358–359, AT V 224)

In his *Replies* to the Sixth set of *Objections*, Descartes admits that we cannot understand how God could have made the “eternal truths” different than we take them to be. But just as in his letters, he insists that God could have done this:

[T]here is no need to ask how God could have brought it about from eternity that it was not true that twice four make eight, and so on; for I admit that this is unintelligible to us. Yet on the other hand I do understand, quite correctly, that there cannot be any class of entity that does not depend on God; I also understand that it would have been easy for God to ordain certain things such that we men cannot understand the possibility of their being otherwise than they are. (CSM II 294, AT VII 436)

Notice that Descartes’s doctrine seems to clash with his own claim, in the Second Set of *Replies*, that at least some propositions can never be doubted at all. For to hold that God could (a) have made it *false* that *not-(p and not-p)* or that $1 + 1 = 2$, but (b) created our minds in such a way that we cannot see those propositions except as being true, seems tantamount to doubting those propositions, at least in some minimal way. Here the consistency of Descartes’s position is under maximum strain.

Most philosophers today, including the present writer, would reject Descartes’s doctrine of the creation of the eternal truths. They would hold, as Leibniz held, that the truths of logic and pure mathematics are true in all the possible worlds that God could have created, so that not even an omnipotent God could dictate or alter them. They would also agree with the view, put forward by Aquinas, that omnipotence does not require the power to do logically impossible things but only the power to

do whatever is logically possible. Nevertheless, Descartes's doctrine should not be dismissed as a mere anomaly or eccentricity in his thought. For the doctrine is in a certain sense deeply Cartesian; it is a striking manifestation of what might be called Descartes's "philosophical radicalism." Other manifestations of this radicalism are, for example, Descartes's carrying traditional skeptical arguments to the extreme of doubting the existence of the entire physical world; his attempt to rebuild all knowledge from one single point of absolute certainty; his sharp dualism of thinking and unextended, versus extended and unthinking, substance; and his view that finite substances depend on God not only for their original creation but also for their continuance in existence from one moment to the next. Descartes's notions that omnipotence entails even the power to do logically impossible things and that God legislated the principles of logic and mathematics bespeak the same radical cast of mind. One need not *agree* with these notions to recognize that they are inherent parts of Descartes's thought.

It might now seem, however, that Descartes's radical doubt of reason depends on his doctrine of the creation of the eternal truths. The idea would be that such propositions as $2 + 2 = 5$ or *both p and not p* are so crazy that they could be true only if God made them true. Such a dependence of the doubt of reason on the doctrine of the creation of the eternal truths would greatly reduce the interest of Descartes's doubt of reason, just because that doctrine is so mind-boggling. So, it is important to see that the radical doubt of reason does *not* really have to depend on the doctrine of the creation of the eternal truths. For if it makes any sense to suppose that God could have made $2 + 2 = 5$ true, then it also makes sense to suppose that $2 + 2 = 5$ could be *unalterably* true; and if it makes any sense to suppose that God could have made *both p and not p* true, then it also makes sense to suppose that *p* and *not p* could both *be* true whether God *made* them true or not. But then the doubt of reason can still be generated: why couldn't $2 + 2 = 5$ be unalterably true, yet God have so made our minds that we can't see the sum of 2 and 2 as being anything other than 4; and why couldn't *both p and not p* be true even without God's having made them so, yet God have so made our minds that we can't see *both p and not p* as being anything but false? To allow the possibility that God gave us radically defective minds is to allow that it could be that (a) 2 plus 2 is unalterably 5, that for some *p*, *p and not p* is unalterably true, that the radii of a circle are unalterably unequal, and so on, and yet that (b) God has so made us that $2 + 2 = 4$ is assent-compelling for us,

not both p and not p is assent-compelling for us, and the radii of a circle are equal is assent-compelling for us, and so forth. Briefly put, the point is that as implausible as Descartes's doctrine of the creation of the eternal truths may be, that doctrine is not necessary to generate his doubt of reason; for suppose that the doctrine is wrong and that not even God could alter the principles of logic and mathematics. Then Descartes can still maintain that an omnipotent God could deceive us about those principles, simply by creating our minds in such a way that the principles that seem necessarily true and indubitable to us are different from the ones that actually and unalterably hold.⁴⁵

4.2.4 *The Validation of Reason*

In light of Descartes's doubt of reason, it appears that the only way to solve the problem of the circle that would be consistent with Descartes's own position would be to show that, appearances to the contrary notwithstanding, it is possible to use reason to overcome a doubt about the reliability of reason itself. Interestingly enough, two contemporary Descartes scholars, Alan Gewirth and Harry Frankfurt, have argued that Descartes undertakes exactly this project and actually succeeds in carrying it out.⁴⁶ Both of these writers offer far-reaching and complex interpretations of Descartes's theory of knowledge, and we cannot do justice to their work here. Rather, we shall briefly outline their defense of Descartes. Then, we shall argue that their defense—or, more precisely, a defense of our own built partly on a key point in their defense—is, indeed, viable.

The key point common to both Gewirth's and Frankfurt's defense of Descartes is that his doubt concerning human reason is not arbitrary or

⁴⁵ I am indebted to James Van Cleve for this point. Margaret Wilson makes the same point in her *Descartes*, pp. 33–34.

⁴⁶ The original source for this approach to the problem of the circle is Alan Gewirth, "The Cartesian Circle," 368–95. Gewirth also expounds his view in two more recent articles: "The Cartesian Circle Reconsidered," 668–85, and "Descartes: Two Disputed Questions," 288–96. All three of these articles are reprinted in Willis Doney ed., *Eternal Truths and the Cartesian Circle* (New York: Garland 1987). The latter two of the articles are part of a very useful exchange with Anthony Kenny, to which Kenny's contribution is his "The Cartesian Circle and the Eternal Truths." Frankfurt's contributions, in which he acknowledges Gewirth's influence, are in his "Descartes's Validation of Reason" and in his *Demons, Dreamers, and Madmen* (esp. chap. 15).

willful. Rather, as Descartes himself says in several places, his doubt is based on reasons. Thus, Gewirth writes:

At the outset of his enterprise, Descartes had laid it down that “it will be sufficient to reject all my opinions, if I find *some reason for doubting* in each one”; at the end of the *First Meditation*, accordingly, the doubt is said to be based upon “valid and meditated reasons,” and in the *Third*, when it is stated in its most extreme form, Descartes calls it a *Metaphysica dubitandi ratio* [a metaphysical reason for doubt] which “depends upon” his “opinion” concerning his deception by an omnipotent God. The very nature of the doubt as conceived by Descartes thus involves an ineluctable rational element: “the must be some *reason* for doubting before one determines oneself thereto.” (Quoting from *Meditation I* [CSM II 12, SPW 76, AT VII 18; CSM II 15, SPW 78–79, AT VII 21–22] and *Meditation III* [CSM II 25, SPW 88, AT VII 36], respectively; and from Descartes’s letter to Clerselier of January 12, 1646 [CSM II 270, AT IXA 204].)⁴⁷

In John Cottingham’s translation, the passage that Gewirth quotes from the letter to Clerselier says, “before we can decide to doubt, we need some reason for doubting” and continues “and that is why in my First Meditation I put forward the principal reasons for doubt.” As one recent commentator puts it, “the First Meditation’s method of universal doubt clearly indicates that doubting, far from being an arbitrary enterprise, requires us to attend to reasons, even such that we may later find to be questionable or false.”⁴⁸ The key point, then, is that the deceiving-God hypothesis is supposed to provide a *reason or argument* for doubting, and that at least by the time we get to *Meditation III*, it is supposed to provide a reason for doubting even the reliability of reason. Therefore, argue Gewirth and Frankfurt, Descartes can refute this reason for doubt (i.e., disqualify it as a reason for doubt) by showing that the deceiving God hypothesis is not a *good reason* for doubting the reliability of reason. But that is exactly what Descartes’s proof of a veracious God shows; for it shows that the deceiving God hypothesis is ruled out by, or is incompatible with, a chain of argument built entirely on clear and distinct perceptions. Notice that this

⁴⁷ Gewirth, “The Cartesian Circle,” p. 388.

⁴⁸ Noa Naaman-Zauderer, *Descartes’ Deontological Turn*, p. 110.

defense of Descartes does not assume that clear and distinct perceptions must be true, which would, of course, beg the question. Rather, it turns on the point that the most careful and conscientious application of reason (i.e., one that involves only clear and distinct perceptions and is not subject to the weaknesses of sense perception, memory, and muddled or careless thinking) leads to the conclusion that the reason Descartes had for doubting the reliability of reason is a bad one, thereby nullifying it as a reason for doubt.

Nevertheless, it may initially seem that this defense of Descartes is open to a fatal objection; for the *content* of Descartes's doubt is precisely whether the best possible use of reason is reliable or trustworthy. If his doubt pertained to anything else—say, only to the reliability of the senses, or of memory, or of other peoples' testimony, or the like—then showing that the reasons given for such a doubt are defeated by the best use of reason could refute the doubt. But, given that Descartes's doubt extends to whether the best use of reason is itself reliable, the logic of the situation may seem to be as follows. Gewirth and Frankfurt's Descartes argues that since the best use of reason leads to the conclusion that reason itself is reliable, it is unreasonable to doubt whether reason is reliable. The critic, however, can reply that Descartes's radical doubt was precisely whether the best use of reason is reliable, so the argument purporting to show that it is reliable cannot be trusted unless one assumes that this doubt has already been dispelled before giving the argument. Therefore, it may seem, the argument cannot refute the radical doubt.

We shall now argue, however, that this powerful-looking objection can be answered. As already noted, the key point in the Gewirth-Frankfurt defense of Descartes is the correct insistence that Descartes's doubt of reason is itself supposed to be based on reasons. Now, what exactly is the reason that generates Descartes's radical doubt? It is, as Descartes says in *Meditation III*, the idea that an omnipotent God could make him go wrong even about the things he perceives with the utmost clarity: "Whenever my preconceived belief in the supreme power of God comes to mind, I cannot but admit that it would be easy for him, if he so desired, to bring it about that I go wrong even in those matters which I think I see utterly clearly with my mind's eye" (CSM II 25, SPW 88, AT VII 36). Descartes's reason for doubting reason, then, is this: Surely an omnipotent God could easily have given a mere creature like poor little me a radically defective intellect—one that goes wrong even in matters that seem utterly clear and evident.

Notice, however, the following crucial point about this reason for doubting reason: it relies on an inference.⁴⁹ This is the inference from

(1) It is possible that there exists an omnipotent God

to

(2) It is possible that there exists a God who makes me go wrong even about the things I perceive clearly and distinctly.

Indeed, it seems that the inference from (1) to (2) must go through at least one intermediate step, namely:

(1a) It is possible that there exists a God who can do anything.
(derived from step 1 and leading to step 2)

Or perhaps the inference should be seen as involving yet another intermediate step, namely,

(1b) It is possible that there exists a God who can make me go wrong even about the things I perceive clearly and distinctly.
(derived from step 1a and leading to step 2)

The inference from (1) to (2), regardless of how exactly it is construed, is essential to Descartes's radical doubt; for without this inference, Descartes's doubt of reason would be merely willful and arbitrary. It would reduce to the bald assertion that he might be mistaken even about what he perceives most clearly. Furthermore, the inference has to bear an enormous weight. It must be solid enough to convince Descartes that it is at least possible that the things he perceives with the utmost clarity could be false. Indeed, as the passages we cited in connection with Descartes's doctrine of the creation of the eternal truths suggest, the inference must bear the weight of convincing him that it could actually be *false* that $2 + 2 = 4$, that *not both p and not-p*, that *a whole is greater than any of its parts*, and so on; for whether or not omnipotence requires the power to *make* these

⁴⁹ In an early version of this section, I attempted, in effect, to deny this point. Roland P. Blum's objection to this denial led me to see the importance of the point. The material that follows is also influenced by Gewirth's "The Cartesian Circle Reconsidered," sec. 6.

principles false, the inference commits one to admitting that they might in any case *be* false. Descartes's insistence that such principles cannot be doubted while one is attending to them, and his claim in the second set of *Replies* that they can never be doubted at all, can serve to impress on us how heavy a weight his inference from God's omnipotence must bear.

But if Descartes can rely—and indeed rely so heavily—on an inference from the possibility that an omnipotent God exists to the possibility that this God deceives him even about what seems utterly clear to him, then why can he not rely on the chain of reasoning which leads to the conclusion that although an omnipotent God who *could* deceive him does indeed exist, this God is also an absolutely perfect being who, therefore, *would not* deceive him about his clear and distinct perceptions? You may object that Descartes's proof of a nondeceiving God is neither as simple nor as obvious as is the inference from God's omnipotence to his possibly deceiving us about even the simplest things. But this objection, even if correct, is here besides the point; for the question posed by the problem of the circle is not whether Descartes's specific argument for God's existence and veracity actually succeeds in dispelling the doubt of reason but whether any such argument *could* succeed. And our suggestion is that it could. For once we grant the legitimacy of the use of reason required to infer the possibility of our going wrong about the simplest things from the possibility that there is an omnipotent God, there is no reason in principle to deny the legitimacy of the use of reason which leads to the conclusion that the omnipotent God who actually exists is a perfect being who, while still fully *able* to deceive his creatures, would not *wish* to do so. Indeed, consistency requires that if we allow the former use of reason to be legitimate, then we must also allow the latter to be legitimate. As James Van Cleve incisively put it: "if reason can be used to attack reason, then it can also be used to defend reason."⁵⁰

This defense of Descartes can also be put another way. If we look at the inference from the possibility that there is an omnipotent God to his possibly deceiving us about the simplest things *in complete isolation* from all other considerations bearing on the existence and nature of God, then we may suppose that we might be deceived about the things we perceive most clearly; for since an omnipotent God would surely have the power to deceive us, nothing, so far, prevents us from suspecting that he might or does deceive us. So, the inference provides a reasoned basis—Descartes's "very slight and, so to speak, metaphysical" reason—for

⁵⁰ In conversation.

doubting the reliability of reason. However, once we have come to a fuller view of the considerations bearing on God's existence and nature and have deployed the chain of reasoning which leads to the conclusion that a God exists who is not only omnipotent but also supremely good, the suspicion is removed and the doubt is dispelled; for although the most careful reasoning we are capable of leads us to believe that God *could*, if he wished, deceive us even about the simplest matters and although this reasoning thereby initially provides some reason for doubting reason itself, the same kind of reasoning, further and more deeply pursued, shows that God does not in fact (wish to) deceive us. Thus, the original reason for doubt can operate only on the condition that we accept the argument that eventually dispels it. As some philosophers might put it, the reason for doubt is only a "*prima facie*" or "defeasible" reason, which can be disposed of by further reflection.

To put it yet one more way: although "a perfect God exists" does entail (2)—that it is possible that there exists a God who makes me go wrong even about the things I perceive clearly and distinctly—it also entails that God does not (wish to) make me go wrong about the things I perceive clearly and distinctly. Thus, Descartes does not show that it is *impossible* that there is a God who deceives him about the things he perceives clearly and distinctly; he does not show that (2) above—that "it is *possible* that there exists a God who makes me go wrong even about the things I perceive clearly and distinctly"—is false. Rather, he shows that it is *false* that there is a God who actually does deceive him even about the things he perceives clearly and distinctly. And this is enough to refute the radical doubt of reason.

It might be asked, Which of the two strategies earlier discussed does our defense of Descartes employ? Are we saying that Descartes can know that whatever he clearly and distinctly perceives is true before knowing that a nondeceiving God exists (the vindication-not-needed strategy) or that he can know that God exists and is no deceiver before knowing that whatever he perceives clearly and distinctly is true (the criterion-not-needed strategy)? The answer is that according to the defense we have proposed (and assuming for the moment that Descartes's theological argument is sound), Descartes comes to know those two propositions *simultaneously*. Before the argument for God's existence and veracity is complete, the inference from God's omnipotence to his possibly deceiving us provides a reason to doubt clear and distinct perceptions. But the moment the argument is completed, that reason is nullified, and Descartes comes to know both that a nondeceiving God exists and that his clear and distinct

perceptions are true.⁵¹ The entire argument is, as it were, “on probation” until its final conclusion is reached. Thus, our defense relates to the two strategies as follows. In common with both strategies, it recognizes that Descartes’s theological argument uses clear and distinct perceptions. But in contrast to both strategies, it does not assume that the perceptions used in the argument are known to be reliable before the argument is complete. Rather, they are used only because they constitute the most careful use of the intellect we are capable of. However, in common with the criterion-not-needed strategy, our defense holds that once the theological argument is complete, reason is validated (since the reason that was given for doubting reason is nullified). And in contrast to the vindication-not-needed strategy, our defense holds that the purpose of the theological argument is not merely to justify appealing to the results of past uses of reason; rather, its purpose is to provide a validation of reason itself.

4.3 A Final Criticism of the Core Argument

If we assume that the problem of the Cartesian Circle can be disposed of (either in the manner we have proposed, or, failing that, by restricting the Cartesian doubt to sense-perception), then Descartes’s *Meditation III* argument for God’s existence may, at this point, still seem worthy of acceptance; for, as we saw in section 4.1.3, the core argument’s first premise—that the cause of an idea must have as much formal reality as the idea contains objective reality—can be reasonably maintained, especially if it is taken as a basic premise, rather than supported by the subargument Descartes himself offers for it. May we conclude, then, that Descartes’s argument is successful?

There is reason to doubt that such an evaluation would be correct. For consider the core argument’s *second* premise, namely, “Only a perfect God has as much formal reality as my idea of God contains objective reality.” This premise means that only a perfect God has as much reality as my idea of God represents him as having—in other words, that my idea of God is so rich in its informational content that only a supremely perfect God himself has as much reality as that content displays. This claim is a basic premise in the argument; it is not derived from any other

⁵¹ Cf. Van Cleve, “Foundationalism, Epistemic Principles, and the Cartesian Circle,” p. 59, n. 14.

premise of the argument and it is supposed to be known just by contemplating one's idea of God. But given the limits of the human mind, is such a claim plausible? In the Fifth set of *Objections*, Gassendi wrote:

But . . . the idea, or its objective reality, is not to be measured by the total formal reality of the thing (i.e., the reality which the thing has in itself) but merely by that part of the thing of which the intellect has acquired knowledge (i.e., by the knowledge that the intellect has of the thing). Thus you will be said to have a perfect idea of a man if you have looked at him carefully and often from all sides; but your idea will be imperfect if you have merely seen him in passing and on one occasion and from one side . . . You claim that there is in the idea of an infinite God more objective reality than in the idea of a finite thing. But first of all, the human intellect is not capable of conceiving infinity, and hence it neither has nor can contemplate any idea representing an infinite thing. Hence if someone calls something 'infinite' he attributes to a thing which he does not grasp a label which he does not understand. (CSM II 199–200, AT VII 285–6; see also CSM II 205–6, AT VII 295)

Gassendi's objection might be put this way. The degree of objective reality contained in an idea depends on the formal reality of only that *part* of the idea's object that we understand, not on the full degree or amount of formal reality possessed by the object itself. [Notice that this goes directly against the second key idea behind the core argument; namely that the degree of reality that an idea represents its object as having (its degree of objective reality) depends (*solely*) on the degree of reality possessed by the object itself (on its object's degree of formal reality)]. But the human mind cannot grasp the infinite degree of formal reality had by God. So Descartes's premise that only a perfect God has as much formal reality as his idea of a perfect God contains objective reality is simply false. Many things that are far less perfect than God have as much formal reality as Descartes's (necessarily inadequate) idea of God contains objective reality.

Replying to a similar objection, which had also been raised in the First set of *Objections*, Descartes wrote:

At this point, however, he shrewdly asks whether I am 'clearly and distinctly aware of the infinite' . . . Let me say first of all that

the infinite, *qua* infinite, can in no way be grasped. But it can still be understood, in so far as we can clearly and distinctly understand that something is such that no limitations can be found in it, and this amounts to understanding clearly that it is infinite. (CSM II 81, AT VII 112)

Replying to Gassendi himself, Descartes makes the same point at greater length:

I shall make one point about the idea of the infinite. This, you say, cannot be a true idea unless I grasp the infinite; you say that I can be said, at most, to know part of the infinite, and a very small part at that, which does not correspond to the infinite any better than a picture of one tiny hair represents the whole person to whom it belongs. My point is that, on the contrary, if I can grasp something, it would be a total contradiction for that which I grasp to be infinite. For the idea of the infinite, if it is to be a true idea, cannot be grasped at all, since the impossibility of being grasped is contained in the formal definition of the infinite. Nonetheless, it is evident that the idea which we have of the infinite does not merely represent one part of it, but does really represent the infinite in its entirety. The manner of representation, however, is the manner appropriate to a human idea; and undoubtedly God, or some other intelligent nature more perfect than a human mind, could have a much more perfect, i.e., more accurate and distinct, idea. Similarly we do not doubt that a novice at geometry has an idea of a whole triangle when he understands that it is a figure bounded by three lines, even though geometers are capable of knowing and recognizing in this idea many more properties belonging to the same triangle, of which the novice is ignorant. Just as it suffices for the possession of an idea of the whole triangle to understand that it is a figure contained within three lines, so it suffices for a true and complete idea of the infinite in its entirety if we understand that it is a thing which is bounded by no limits. (CSM II 253–4, AT VII 367–8)

Again, in one of his letters, Descartes says:

[I]t is possible to know that God is infinite and all-powerful although our soul, being finite, cannot comprehend or conceive Him. In the same way we can touch a mountain with our hands but we cannot put our arms around it as we could put them around a tree or something else not too large for them. To comprehend something is to embrace it in one's thought; to know something it is sufficient to touch it with one's thought. (To Mersenne May 27, 1630, CSMK 25, AT I 152)

In these passages, Descartes can be seen as answering in reply to Gassendi's objection, in effect, "In order for the core argument's second premise to be true, I need not have a completely adequate understanding of God's nature. I need only know that his power, goodness, and knowledge have no limits."

In his *Descartes*, Bernard Williams interprets this reply as follows: We can clearly and distinctly conceive that God is infinite, but not *how* he is.⁵² Williams then argues that this reply is unsuccessful:

Descartes's course . . . is, in effect, that he can clearly and distinctly conceive that God is actually infinite, but not *how* he is. But that this is an unsatisfactory line of defence can be seen if one reverts to Descartes's own helpful analogy of the man who had the idea of the very complex machine. From the fact that a man has this idea . . . it could be inferred that either he had seen such a machine (or . . . had been told about it) or that he was clever enough to invent it. But clearly such inferences will hold only if the man has a quite determinate idea of the machine. If a man comes up and says that he has an idea of a marvelous machine which will feed the hungry by making proteins out of sand, I shall be impressed neither by his experience nor by his powers of invention if it turns out that that is all there is to the idea, and that he has no conception, or only the haziest conception, of how such a machine might work.⁵³

Williams is not here questioning Descartes's principle that the cause of an idea must have as much formal reality as the idea contains objective

⁵² Williams, *Descartes: The Project of Pure Enquiry*, p. 144.

⁵³ Williams, *Descartes: The Project of Pure Enquiry*, pp. 144–45.

reality; he allows that one could use such a principle to infer a very impressive cause from a very impressive idea. Rather, Williams is questioning whether, given this principle, our idea of God is sufficiently impressive to warrant the inference to God as its cause. His point is that since, for example, we do not understand how God's omnipotence works, or how he can possess knowledge of all past, present and future events, or how he can allow evil if he is supremely good, our idea of God is not sufficiently rich or detailed in its content to warrant an inference to God as its cause. For it is simply not the case that only a perfect (omnipotent, omniscient, omnibenevolent) God has as much formal reality as our idea of him contains objective reality.

To see Williams's point more clearly, suppose that the content of a man's idea of a machine that can turn sand into protein were just this: "machine that can turn sand into protein." Is it true that only a machine that really can turn sand into protein would have as much complexity ("formal reality") as does that idea's content ("objective reality")? No, for such a machine would have to have many extraordinary components, capable of executing many extraordinary chemical reactions, that are not represented in this man's idea. So it's false that only a machine that can really turn sand into protein has as much complexity as does the content of this man's idea. Likewise, suppose that the content of a human's idea of God is just: "being who is all-powerful, all-knowing, and supremely good." Is it true that only a being that really has these properties has as much formal reality as that idea's content or "objective reality"? No, because such a being would have to have many properties (including detailed knowledge of every past, present, and future truth, knowledge of how to create something out of nothing, knowledge of how apparent evils all lead to a greater good) that are not represented in a human idea. As Daniel Dennett neatly says in a recent article, Descartes seems to be "confusing *an idea of a wonderful thing* with a *wonderful idea of a wonderful thing*."⁵⁴ It seems just false, then, that only a perfect or infinite God has as much formal reality as a human being's idea of God has objective reality.

This rebuttal of Descartes's reply to Gassendi seems very powerful. In the end, then, it is doubtful that Descartes's *Meditation III* argument, for all its profundity, provides a compelling case for the existence of God. As we shall see in chapter 5, however, Descartes offers us yet another proof of God's existence—the famous Ontological Argument.

⁵⁴ Daniel C. Dennett, "Descartes's Argument from Design," 333–345, p. 338.

Meditation IV

Error, Freedom, and Evil

1. The Issues of the *Fourth Meditation*

In his *Fourth Meditation*, Descartes addresses two questions that naturally arise from his previous *Meditations*: (1) How can I make errors if God is not a deceiver? (2) How can God be a supremely perfect being if he has so created me that I am capable of making errors? The first problem must be addressed for the Cartesian doubt to be completely and permanently cured, for so long as we do not know what the cause of our errors is, there is room for a residual doubt about the reliability of our faculties. The second problem is a special case of the classic “problem of evil”: how can we reconcile the belief in a perfect God with the existence of evil (in this case the existence of error, seen as a particular kind of evil) in the world that he created? The two problems are intertwined in Descartes’s exposition, but they will be discussed separately in this chapter.

2. Error and the Will

Near the beginning of the *Fourth Meditation*, Descartes writes:

I know by experience that there is in me a faculty of judgement which, like everything else which is in me, I certainly received from God. And since God does not wish to deceive me, he surely did not give me the kind of faculty which would ever enable me to go wrong while using it correctly. (CSM II 38, SPW 99, AT VII 53–54)

In other words, since God is not a deceiver, it must be that I cannot go wrong provided I use the intellectual powers that he has given me correctly. So my errors must come from some misuse of my faculties. What then is this misuse? In a nutshell, Descartes's answer is that he has two mental faculties, the intellect and the will, and that error does not come from either one of them separately, but from the joint action of the two.

The intellect is the capacity to have clear and distinct perceptions. Descartes has two important doctrines about it. First, the intellect cannot affirm or deny anything; or, putting it more carefully, by using only the intellect, a person or mind cannot affirm or deny anything. The reason is that the intellect, which Descartes also calls the "faculty of understanding" (CSM II 39, SPW 101, AT VII 57) and "my understanding [which] comes from God" (CSM II 40, SPW 102, AT VII 58), is only the capacity to *understand* a proposition or an idea; it does not enable one to affirm or to deny any proposition, or to affirm or deny that any of its ideas correspond to reality. It immediately follows from this that the intellect alone cannot be the source of error, since error consists precisely in affirming a false proposition or in denying a true proposition (or, Descartes says, in pursuing evil or sin rather than good).¹ Thus, Descartes can safely hold that the intellect is in no way defective or faulty. Second, on the other hand, the (human) intellect is finite or limited; it cannot grasp everything or understand every question; there are some matters that it cannot perceive clearly and distinctly. Descartes insists that the finitude of the human intellect, *per se*, in no way diminishes God's perfection. As he says,

I have no cause for complaint on the grounds that the power of understanding or the natural light which God gave me is no greater than it is; for it is in the nature of a finite intellect to lack understanding of many things, and it is in the nature of a created intellect to be finite. Indeed, I have reason to give thanks to him who has never owed me anything for the great bounty that he has shown me, rather than thinking myself deprived or robbed of any gifts he did not bestow. (CSM II 42, SPW 103, AT VII 60)

¹ Some commentators hold that for Descartes, even affirming a true proposition or denying a false one counts as an error, provided that the affirmation or denial is not based on clear and distinct perceptions. See for example Noa Naaman-Zauderer, *Descartes' Deontological Turn*, pp. 64, 68, 70–71, 78, 85.

The will, by contrast, is the capacity to affirm or to deny (or more generally, to do X or not to do X, where X includes not only affirming or denying, but also pursuing or avoiding); it is thus necessarily involved in assenting to any proposition or in making any judgment. Descartes's most important doctrine about the will is that, unlike the intellect, its scope is unlimited. In the *Fourth Meditation*, he says that "it is only the will, or freedom of choice, which I experience within me to be so great that the idea of any greater faculty is beyond my grasp; so much so that it is above all in virtue of the will that I understand myself to bear in some way the image and likeness of God" (CSM II 40, SPW 101, AT VII 57), and in the *Principles of Philosophy* he even says that "[the will] can in a certain sense be called infinite" (CSM I 204, SPW 171, AT VIIIA 18). What this doctrine means is that there is no proposition grasped by the intellect that the will can neither affirm nor deny, and no course of action that it cannot choose. However, since clear and distinct perceptions are assent-compelling, it does not mean that for any proposition, we can always either affirm or deny it: when we clearly and distinctly perceive a proposition, we can only affirm it. Nor does it mean that we can actually carry out any course of action that we choose, since unlike God we lack the power to do many things.²

It follows that one can choose to affirm or to deny even propositions that surpass one's intellect or understanding—that is, that one does not clearly and distinctly perceive—and *this is precisely how error arises*. As Descartes puts it,

When I look more closely into myself and inquire into the nature of my errors . . . , I notice that they depend on two concurrent causes, namely on the faculty of knowledge which is in me, and on the faculty of choice or freedom of the will; that is, they depend on both the intellect and the will simultaneously . . . [For] the scope of the will is wider than that of the intellect; but instead of restricting it within the same limits, I extend its use to matters which I do not understand. Since the will is indifferent in such cases, it easily turns aside from what is true and good, and this is the source of my error and sin. (CSM II 39–41, SPW 101–102, AT VII 56–58)

² It is not clear whether Descartes means to include, among actions that we can "choose," ones that no human being can actually perform, such as jumping ten meters high.

The corollary of this analysis of error, as the will's affirming or denying propositions that the intellect does not clearly and distinctly perceive, is that if one restricts one's judgments of truth or falsity (and, Descartes says, of goodness and badness) to what one clearly and distinctly perceives, then one cannot go wrong—one is then infallible. Descartes fully accepts this result:

If, whenever I have to make a judgment, I restrain my will so that it extends to what the intellect clearly and distinctly reveals, and no further, then it is quite impossible for me to go wrong. (CSM II 43, SPW 104–105, AT VII 62)

3. Some Possible Objections

Although the secondary literature on Descartes's *Fourth Meditation* is much smaller than that on any of the other five, his theory of error raises fascinating issues. We shall confine our discussion to two possible objections to the theory.

3.1 Assenting and Deciding to Believe

The most common objection to Descartes's theory of error is that by attributing affirmation and denial to the will, it wrongly implies that people can simply decide what to believe, even if there is no basis at all for their decision. E. M. Curley puts the objection this way:

Consider some proposition for which you have literally no evidence at all one way or the other. It may not be as easy as you might suppose to think of an example that strictly satisfies that condition, but I am in that situation with respect to the proposition “it rained three hours ago on Jupiter.” Now, paying careful attention to what is happening in your mind, believe it. Or, if you prefer, disbelieve it. Or do both, in turn. Did anything happen? Unless your experience is very unlike mine, I suspect not. Indeed. I fear that if my salvation depended on my either believing or disbelieving this particular proposition, I should be damned.³

³ E. M. Curley, “Descartes, Spinoza and the Ethics of Belief,” p. 178.

This objection seems to be directed against a view that Descartes never affirms and that he is not logically committed to. Nowhere does Descartes say that we can simply decide what to believe—that (a) for any proposition p that I understand, I can just decide to believe or to disbelieve p . Rather, he holds the different view that (b) for any proposition p that I understand (with the exception of the negations of propositions that I currently perceive clearly and distinctly) there is *nothing to stop me* from believing or assenting to p . This is perfectly compatible with holding, as indeed there is good reason to think Descartes did hold, that humans do not adopt a belief unless there is some reason that moves them to do so—that the will does not affirm a proposition p in a perfectly arbitrary way, without any reason whatsoever. That reason might be a poor or frivolous one; for example, I might believe that it rained on Jupiter because my mother told me so, or because the earth and Jupiter are aligned in a certain way, or because I wishfully think that there are lush forests on Jupiter. But that is not to say that I believe it arbitrarily, and whether humans ever do believe things that way seems to be an empirical question.

Basically, Descartes is only making the common-sense point that we are capable of affirming and denying things before making a clear and reasoned judgment. He is also saying that if we affirm or deny something before making a careful and reasoned judgment, then we are misusing our freedom and that this misuse is the source of error. More grandly, since he holds that there is no proposition we understand that we can neither affirm nor deny, and calls the capacity to affirm or deny a proposition “the will,” he says that the scope of the will is “infinite.” But there is nothing in his texts to indicate that by this he means that we can just decide to believe any proposition. At most, Descartes might be criticized for not saying anything to forestall such a misunderstanding of his position, especially given his talk of infinity.

This is not to say that there are no issues about the extent to which we can control our beliefs. On the contrary, the topic of “doxastic voluntarism” is a controversial one, on which there is a substantial body of literature that addresses questions like, Can we believe something in the face of overwhelming evidence against it? Is wishful thinking, or indoctrination, or superstition, potentially sufficient to neutralize the weight of any such evidence in our minds, as Descartes would presumably say? But it is not necessary to pursue these questions in order to exonerate Descartes from the common criticism that on his view we can just arbitrarily decide to believe anything at all.

3.2 Irresistibility and Freedom

As John Cottingham observes, “the central doctrine of Descartes’ Fourth Meditation is what may be called the ‘doctrine of the irresistibility of the natural light’: when the intellect is confronted with a clear and distinct perception, the will is immediately and spontaneously impelled to give its assent to the truth of the relevant proposition.”⁴ Indeed, we have seen that Descartes holds that clear and distinct perceptions are irresistible or assent-compelling, that is, that during the time that one clearly and distinctly perceives some proposition *p*, one cannot help believing *p*, one cannot doubt that *p*. In terms of his theory about the role of the intellect and the will in making judgments, his position is that as long as the intellect clearly and distinctly perceives *p*, the will cannot do otherwise than assent to *p*. This doctrine is one that he insists upon in many places, for example:

My nature is such that so long as I perceive something very clearly and distinctly I cannot but believe that it is true. (CSM II 48, AT VII 69, SPW 109)

I have already demonstrated that everything of which I am clearly aware is true. And even if I had not demonstrated this, the nature of my mind is such that I cannot but assent to these things, at least so long as I clearly perceive them. (CSM II 45, SPW 106, AT VII 65)

So long as I attend to the proof [that the three angles of a triangle are equal to two right angles], I cannot but believe this to be true. (CSM II 48, SPW 109, AT VII 69)

So long as we attend to a truth which we perceive very clearly, we cannot doubt it. (CSM II 309, AT VII 460)

On the other hand, Descartes is equally insistent that when we assent to a clear and distinct perception, we do so *freely*; indeed he thinks that we are then at our freest—that assenting to a clear and distinct perception is the “best” kind of freedom that we can have. When one brings to mind the historical context—Descartes’s time was one when following the dictates of reason or the lead of the evidence rather than accepting the dogmas sanctioned by authority was neither the favored method of

⁴ John Cottingham, *A Descartes Dictionary*, pp. 64–65.

fixing belief nor conducive to one's personal safety—Descartes's position may well strike us as a salutary celebration of intellectual freedom. But from a strictly philosophical point of view, his position gives rise to what looks like a powerful objection: how can it possibly be true that we are free when we assent to a clear and distinct perception, if such a perception compels our assent—if we cannot do otherwise than to assent?

To see how Descartes could deal with this problem, we need to investigate more closely what he thinks freedom is. The text that scholars frequently cite as most important in this regard is this passage from *Meditation IV*:

The will simply consists in our ability to do or not do something (that is, to affirm or deny, to pursue or to avoid); or rather, it consists simply in the fact that when the intellect puts something forward for affirmation or denial or for pursuit or avoidance, our inclinations are such that we do not feel we are determined by any external force. (CSM II 40, AT VII 57, SPW 101–102)

At least since the appearance of an important article by Anthony Kenny, scholars have seen this passage as problematic.⁵ Its first part (up to the semicolon) looks like a straightforward affirmation of what C. P. Ragland, slightly adapting a term coined by Harry Frankfurt, calls “the principle of alternative possibilities”—the principle that X is a free act only if the agent has the ability to either do X or not do X, or the ability to “do otherwise.”⁶ This principle states a condition for freedom that is highly intuitive and that many philosophers hold to be the sine qua non of freedom. Another place where Descartes seems to endorse the principle, and indeed to apply it directly to acts of assenting to true propositions, is his *Principles of Philosophy*, Part One, section 37:

We do not praise automatons for accurately producing all the movements they were designed to perform, because the production of these movements occurs necessarily. It is the designer

⁵ Anthony Kenny, “Descartes on the Will,” pp. 1–31.

⁶ C. P. Ragland, “Alternative Possibilities in Descartes’s Fourth Meditation,” 379–400 and “Descartes on the Principle of Alternative Possibilities,” 377–94. I say “slightly adapting” because Frankfurt treats the principle as a (proposed) necessary condition for moral responsibility, whereas Ragland treats it as a necessary condition for freedom.

who is praised for constructing such carefully-made devices; for in constructing them he acted not out of necessity but freely. By the same principle, when we embrace the truth, our doing so voluntarily is much more to our credit *than would be the case if we could not do otherwise.* (CSM I 205, SPW 172, AT VIIIA 18–19 [emphasis added])

In the second part of the above passage from *Meditation IV*, however, Descartes gives what looks like a quite different definition of freedom. First, he brings in his view that any free act (of affirming or denying, seeking or avoiding) must follow upon some perception provided by the intellect. Second, he seems to be saying that what makes the act free is that we do not feel compelled to perform it by some external agent or force or by any factor other than ourselves.

This second formulation gives rise to two questions: (1) Does Descartes really mean that in order to act freely, we need only *feel* that our act is spontaneous or undetermined by any external force, or does he mean that our act must actually *be* undetermined or spontaneous? (2) How does the second formulation (about not being determined by any external force) relate to the first one (about having the two-way power of choice, or being able to do otherwise)? Does it, as some have suggested, retract it, or does it merely explicate it in some way? With respect to question (1), we shall here simply follow the Israeli scholar Noa Zaaman-Zauderer, who says, in her recent fine book on Descartes: “Descartes does not mean that our freedom demands only that we feel ourselves undetermined by external coercion, but also that we *be* so undetermined.”⁷

With respect to (2), matters are more complicated. The second formulation spells out something that must not be the case in order for the first formulation to apply: we must not be determined in our choice by any external force. But the second formulation is certainly not equivalent to the first, for one could be undetermined to do X by any external force and yet unable to do anything other than X. This would be the case if one were determined (compelled) to do X by some factor *internal* to oneself, for example by an addiction, or a compulsion, or some mental illness. Indeed, there are philosophers who believe that no human actions are free

⁷ Noa Naaman-Zauderer, *Descartes' Deontological Turn*, p. 103 n6. Naaman-Zauderer cites J. K. Campbell, C. P. Ragland, and Vere Chappell as commentators who “rightly emphasize” this point.

because they are all caused by internal factors that are links in a causal chain of events that ultimately extends to events that occurred before one was born and that were therefore out of one's control. This "hard determinist" position, as it is called, rests on the twin assumptions that (a) all events including human actions are caused by antecedent events, and (b) an event's being caused by antecedent events is incompatible with its being a free action. We shall not delve into the complex issue of whether (a) and (b) are both true, since that would require a full-scale treatment of the traditional philosophical problem of freedom and determinism that we cannot undertake here. But it is clear that Descartes was not a hard determinist, so that he would have rejected (a) or (b) or both (a) and (b); for he holds that the will is free *by definition*, which entails that acts of will, or volitions, are free. As Vere Chappell says,

It is [Descartes's] view that every volition is free, and that it is so, furthermore, of necessity. For it is the essence of will, as he puts it, to act freely; willing is free by nature. . . . It is not that men have the power of willing, some of whose exercises are free and some not. Rather they just have the power of free-willing, or willing-freely.⁸

It follows that an action that results from a volition or an act of will—one that is voluntary—is free; an action can fail to be free only by not resulting from a volition, by not being voluntary. But Descartes certainly holds that there are voluntary actions, so he is logically committed to holding that there are free actions.

Descartes also holds, however, that freedom is not an all-or-nothing matter; rather, according to him there are degrees or grades of freedom. These degrees are based on the extent to which a choice or action is based on *good reasons*. The lowest grade of freedom, according to this criterion, attaches to actions that result from choices that are, from a rational point of view, arbitrary—ones that are made even though there are no good reasons for preferring any of the possible alternatives, or when the reasons for choosing among them are evenly balanced; Descartes calls this grade of freedom "freedom (liberty) of indifference."⁹ An action that

⁸ Vere Chappell, "Descartes's Compatibilism," p. 179. Chappell cites CSM II 117/AT VII 166, AT XI 359 and CSM I/AT 343, among others, as supporting passages.

⁹ He thus gives this term a different meaning than did the Scholastics, who used it simply to mean the two-way power to perform or not to perform an action.

was motivated only by impulse, passion, or emotion would presumably exhibit this lowest grade of freedom. Intermediate grades of freedom would attach to choices that are based partly on good reasons; as Descartes puts it in a letter of May 2, 1644 to Mesland: "I did not say that a person was indifferent only if he lacked knowledge, but rather, that he is more indifferent the fewer reasons he knows which impel him to choose one side rather than another" (CSMK 233, AT IV 115). The highest grade of freedom attaches to actions (affirmations and denials, pursuits and avoidances) that result from a clear and distinct perception of what is true or what is good; Descartes calls this grade of freedom "freedom of spontaneity."

This reference to freedom of spontaneity brings us back to the problem of how assenting to a clearly and distinctly perceived proposition can be a free act if we cannot but assent to it—if such a proposition is assent-compelling. It is clear that Descartes holds that in such cases we are free, indeed maximally free, for he writes:

During these past few days I have been asking whether anything in the world exists, and I have realized that from the very fact of my raising this question it follows quite evidently that I exist. *I could not but judge that something which I understood so clearly was true;* but this was not because I was compelled so to judge by any external force, but because a great light in the intellect was followed by a great inclination in the will, and *thus the spontaneity and freedom of my belief was all the greater in proportion to my lack of indifference.* (CSM II 41, AT VII 58–59, SPW 102; emphasis added)

Here it seems that Descartes is saying that when I clearly and distinctly perceive the *cogito* (or, by implication, any clear and distinct proposition), (a) I cannot do otherwise than assent to it, but (b) this assent is perfectly and maximally free. Our problem is: how can this possibly be true? Isn't it just incoherent?

4. The Coherence of Cartesian Freedom

Descartes could give at least two different responses. To understand the first response, it may be helpful to first consider first what he thinks about choosing good, as opposed to evil or bad, things and actions. His

position is that when one clearly and distinctly perceives X to be better than Y, one has no choice but to select X. He rejects the possibility of what philosophers call “weakness of will” (*akrasia*), which can be defined as “choosing something that one *knows* to be bad.” As he puts it:

Since our will works so as to pursue or avoid only what our intellect represents as good or bad, we need only judge well in order to act well, and to judge as well as we can to do our best. (CSM I 125, SPW 34, AT VI 28)¹⁰

If I always saw clearly what is true and good, I should never have to deliberate about the right judgment or choice; in that case, although I should be wholly free, it would be impossible for me ever to be in a state of indifference. (CSM II 40, SPW 102, AT VII 58)

It seems to me certain that a great light in the intellect is followed by a great inclination of the will; so that if we see very clearly that a thing is good for us, it is very difficult—and, on my view, impossible, as long as one continues in the same thought—to stop the course of our desire. (CSMK 233, AT IV 116)

If we saw it [the sinfulness of some contemplated action] clearly, it would be impossible for us to sin, as long as we saw it in that fashion; that is why they say that whoever sins does so in ignorance. (CSMK 234, AT IV 117)

Notice how, in the second passage, Descartes treats a clearly known truth in the same way as a clearly known good—a point to which we shall return shortly.

Descartes’s view that when X is clearly known to be better than Y or, more simply put, when X is “good,” one cannot but choose X, has a noble

¹⁰ This passage is from Descartes’ *Discourse on the Method*, originally written in French. I have altered Cottingham’s translation a little. Descartes’s highly idiomatic text says: “d’autant que notre volonté ne se portant à suivre ni à fuir aucune chose, que selon que notre entendement la lui représente bonne ou mauvaise, il suffit de bien juger, pour bien faire, et de juger le mieux qu’on puisse, pour faire aussi tout son mieux.” Cottingham translates “ne se portant à suivre ni à fuir” as “tends to pursue or avoid.” That is too weak, insofar as “tends” suggests or implies a mere tendency rather than an invariable way of operating, which is why I have used the locution “works in such a way as to pursue or avoid.” A completely literal translation of Descartes’s words would be, “(Inasmuch as) our will carrying (carries) itself so as neither to pursue nor to avoid anything except what our intellect represents to it as good or bad, it suffices to judge well in order to act well, and to judge as well as one can to also do one’s best.”

lineage in the thought of Plato and Aristotle. The thought that if one really *knows* that X is good—that is, that X is the best of the available alternatives—one cannot fail to pursue X, is a powerful one. Descartes would support it by holding that when one pursues the bad rather than the good, this can only be because one is prevented from clearly and distinctly perceiving the good by inattention, emotion, or passion, or the like—because, as the saying goes, one is “blinded” by these. But on the other hand, the thought that a person may fail to pursue something that she knows to be best—that is, the thought that there is such a thing as weakness of the will—seems to be borne out by a common experience, namely, that of seeing people who we think “know better” make bad choices. We shall not delve into the controversial question of whether or not there is such a thing as weakness of the will. Rather, the important point for our purposes is that even if knowing that X is good inevitably leads to choosing X—that is, even if there is no such thing as weakness of will—this does not mean that the choice of X was unfree. The mere knowledge that X is good obviously has no bearing on freedom or lack thereof. But adding that this knowledge inevitably leads us to choose X (whether this addition is correct or not) does not seem to render the choice of X unfree. To take a parallel case, most epistemologists hold that “S knows that *p*” entails “S believes that *p*”; it does not follow that when someone knows that *p*, her belief that *p* or assent to *p* is unfree, forced, or coerced.

This point bears on the issue of whether the assent-compellingness of clear and distinct perceptions is compatible with free assent to them. For as Margaret Wilson notes, there is in Descartes’s thought an “assimilation of truth and goodness.”¹¹ This assimilation comes out in Descartes’s expressions, “reasons of truth and goodness,” “what was true and good” (CSM II 40, SPW 102, AT VII 58), “what is true and good” (CSM II 41, SPW 102, AT VII 58), “wholly true and good” (CSM II 42, SPW 104, AT VII 60). Thus, just as he thinks that the fact that a clear and distinct perception of the good impels the will to choose the good does not entail that the choice is unfree, so he thinks that the fact that a clear and distinct perception of a true proposition is assent-compelling does not entail that the assent is unfree. Even apart from the assimilation of truth to goodness, this latter point seems right. Suppose that you are faced with a choice between assenting to $1 + 1 = 1$, $1 + 1 = 2$, and $1 + 1 = 3$, and that at time t

¹¹ Margaret Wilson, *Descartes*, p. 143.

you clearly and distinctly perceive $1 + 1 = 2$. Then do you really have any alternative at time t to assenting to that proposition—can you really “do otherwise?” It seems not. Yet it also seems wrong to say that your assent is unfree; on the contrary it seems to be, as Descartes would say, completely spontaneous—a kind of effortless acquiescence with the obvious.

This arithmetical example gives rise, however, to a possible objection that we should address before turning to the second response that Descartes could give to the “incoherence” charge. Descartes considers one’s grasp or understanding of $1 + 1 = 2$ to be an act of the intellect, and one’s assent to $1 + 1 = 2$ to be an act of the will. But, as a number of commentators have said, anyone who understands what a simple mathematical proposition like $1 + 1 = 2$ means thereby also sees that it is true, so there does not seem to be any role for a distinct act of the will here. Bernard Williams puts the point this way:

Clearly and distinctly to understand the proposition about the angles of a triangle [that they must equal two right angles], for example, is to see that it is necessarily true. But once this step has been taken . . . the theory of assent is in difficulty. For if in this sense I clearly and distinctly understand a proposition—that is to say, I can see that it is true—there is nothing else I have to do in order to believe it: I already believe it. The will has nothing to do which the understanding has not already done.¹²

To answer this objection on Descartes’s behalf, we need to start with a point about judgment.¹³ A judgment has two distinguishable components: a propositional content and a propositional attitude toward that content. For example, in the judgment, “the cat is hiding under the table,” the propositional content is *the cat’s hiding under the table*, and the propositional attitude is that of affirming this content. Now for Descartes, the intellect is nothing but the capacity to grasp propositional content (and also mere concepts like *cat*), and the will is simply the capacity to affirm or deny propositional content (and also to pursue or avoid perceived “goods” and “bads”). Since we do in fact grasp propositional contents as well as affirm or deny them and pursue or avoid things, we do

¹² Bernard Williams, *Descartes: The Project of Pure Enquiry*, p. 183.

¹³ The analysis of the role of intellect and will in judgment that follows is due to David M. Rosenthal, “Will and the Theory of Judgment,” pp. 405–34.

have those capacities, and this is arguably all that Descartes means by saying that we have the “faculties” of intellect and will. Now in the case of a clearly and distinctly perceived proposition, it is not the case that one of those capacities—the capacity to affirm or deny—ceases to function, since we do affirm the proposition. Nor is it the case, of course, that the other capacity, of understanding the proposition, ceases to function, since we cannot affirm or deny something that we do not understand. Rather, the activation of the capacity to understand the content is inevitably accompanied by the activation of the capacity to affirm it. Consider an analogy. Suppose that Henri is a fluent French speaker with normal and currently unimpeded hearing and intelligence. Then if Henri distinctly hears the spoken words “*le chat se cache sous la table*,” he also immediately and invariably understands what is said—that the cat is hiding under the table. But it does not follow that his capacity to hear and his capacity to understand are identical, or that only one of them is operating, or that one of them is superfluous, or anything of the sort. It will not do to respond that there are no possible worlds in which one understands $1 + 1 = 2$ but withholds or denies it, for there are also no possible worlds in which a fluent French speaker with currently unimpeded, normal hearing and intelligence distinctly hears “*le chat se cache sous la table*” but fails to understand what is being said.

Williams’s objection seems to turn on a misleading personification of the intellect and the will that Descartes’s language (and even common language) encourages. Let us explain. The act of assenting to a clear and distinct perception is, from a phenomenological (experiential) point of view, a single act: we do not have an experience of understanding $1 + 1 = 2$, and another, distinct experience of assenting to it. Further, this single act calls for only one agent; indeed if there were two agents involved—an “understander” and an “assenter”—then there would no assent to the proposition that $1 + 1 = 2$. Now obviously understanding must be involved, since we cannot affirm something that we do not understand. So it may seem, just as Williams thinks, that any act of the will would be superfluous. But this line of thinking wrongly treats the understanding and the will each as agents or “actors.” If they were, then only one such agent would be needed and, since clearly the understanding is needed, the will would be otiose. The truth of the matter, however, is that the agent is a person or a mind, and that the will and the understanding are not agents but capacities of that person. There is nothing suspect in saying that when a person assents to a proposition, both of

these capacities are activated, and that in the special case of a clearly and distinctly perceived proposition, the activation of the one necessitates the activation of the other.

We can now turn to the second response that Descartes could give to the charge that he cannot coherently hold that assenting to a clearly and distinctly perceived proposition can be a free act if we cannot do otherwise. This response turns on the interpretation of the principle of alternative possibilities. It is commonly assumed that this principle means that for an act X performed at a time t to be free, it must be possible for the agent to refrain from performing X at that very time, t . This entails that for the act of assenting to a clearly and distinctly perceived proposition p a time t to be free, it must be possible for the assenter to deny or withhold p at that very time, t . To be sure, some commentators have rightly emphasized that it is *only* at the moment of illumination—only at the moment that one clearly and distinctly perceives p —that one cannot but assent to p , and they also seem to think that this somehow mitigates the incoherence of holding both that such assent is free and that it violates the principle of alternative possibilities.¹⁴ But, so far as this writer is aware, no commentator has gone so far as to suggest that, at least in Descartes's eyes, the principle of alternative possibilities is *still satisfied* in such a case. Instead, the fact that at the moment of illumination one cannot withhold or deny p is seen as showing that Descartes is not fully committed to that principle, since the principle is supposed to mean that assent to a clearly and distinctly perceived proposition can be free only if one can deny or withhold the proposition at the very moment of illumination. But it is not necessary to place such a strict interpretation on the principle, nor does it seem fair to Descartes. For Descartes agrees, in fact he asserts, that a proposition that could be clearly and distinctly perceived can be denied because of a lack of concentration, inattention, confusion, or any other factor that prevents one from perceiving it clearly and distinctly. On any such occasion, one *can* do otherwise than to assent to the proposition. Descartes claims only that at the moment that one clearly and distinctly perceives a proposition, one is unable to refrain from accepting it as true. He also says in several places that it is difficult to keep one's focus on a clearly and distinctly perceived truth, and that the moment that one's focus shifts or that one is distracted, one is no longer impelled to assent to the proposition, and can even doubt its truth.

¹⁴ See, for example, Cottingham, *A Descartes Dictionary*, pp. 65–66.

(CSM II 43, AT VII 62; CSM II 48, AT VII 69; CSMK 233–234, AT IV 116). It seems perverse, then, to insist that for assent to a proposition to conform to the principle of alternative possibilities, one must be able to reject or withhold that proposition even when one perceives it clearly and distinctly. Furthermore, as Descartes emphasizes, arriving at a clear and distinct perception often requires careful and sustained effort and thought, in which one freely focuses one's attention on reasons and evidence, and further exercises one's freedom in deliberating upon them, as exemplified in the *Meditations* themselves. Is it reasonable to say, as the usual interpretation of the principle of alternative possibilities would make us say, that the fruit of such efforts—namely, a clear and distinct perception in which “a great light in the intellect [is] followed by a great inclination in the will”—is an unfree or coerced assent to a truth?

Consider an analogy. Suppose that a skilled diver is poised at the end of a diving board. After much training and practice, after taking position and carefully rehearsing the dive in her mind, she dives. Of course, from that point on she has no alternative but to “fall” into the water. Does this mean that her dive, or anything that she did before, during, or after her dive, was unfree or coerced? It seems not; for she had the alternative of refraining from diving, which is what matters to her freedom. Now suppose that, say, a mathematician works long and hard to arrive at the proof of a theorem, discovers the proof, and then keeps the whole proof in mind and perceives that the theorem follows from the premises so clearly and distinctly that he cannot but assent to the theorem. Does this mean that this assent is unfree? It seems not. At most, someone might complain that calling the assent an act is misleading, since it suggests that there is some act of assenting to the theorem that occurs later in time than grasping the proof. But it is clear that the mathematician does assent to or mentally affirm the theorem, and one may be hard pressed to improve on the Cartesian language that calls this assent an act of the will.

We propose, then, the following interpretation of Descartes's position: when Descartes says that our assent to a clear and distinct perception is free, he intends to *include* the thought that we can do otherwise than assent to it: he fully *accepts* the principle of alternative possibilities. But for him that principle entails only that we can withhold or deny the relevant proposition *except* at the very time when are perceiving it clearly and distinctly. At all other times, such as when we are working toward attaining a clear and distinct perception of it, or about to perceive it clearly and distinctly, or no longer perceiving it clearly and distinctly, or

incapable of perceiving it clearly and distinctly, and so forth, we can certainly withhold or deny it, and for Descartes that suffices to satisfy the principle of alternative possibilities. One might call this the “longitudinal,” as opposed to the “cross-sectional,” interpretation of the principle of alternative possibilities.

5. Descartes's Troubling Letter to Mesland

There is a letter in Descartes's correspondence that may seem to clash with our interpretation. This is his much-discussed letter of February 9, 1645 to the Jesuit priest, Denis Mesland, where he says:

“Indifference” . . . seems to me strictly to mean that state of the will when it is not impelled one way rather than another by any perception of truth or goodness. This is the sense in which I took it when I said that the lowest degree of freedom is that by which we determine ourselves to things to which we are indifferent. But perhaps others mean by “indifference” a positive faculty of determining oneself to one or other of two contraries, that is to say, to pursue or avoid, to affirm or deny. I do not deny that the will has this positive faculty. Indeed, I think that it has it not only with respect to those actions to which it is not pushed by any evident reasons on one side rather than on the other, but also with respect to all other actions; so that when a very evident reason moves us in one direction, although morally speaking we can hardly move in the contrary direction, absolutely speaking we can. For it is always open to us to hold back from pursuing a clearly known good, or from admitting a clearly perceived truth, provided we consider it a good thing to demonstrate the freedom of our will by so doing. (CSMK 245, AT IV 173)¹⁵

At least four things seem to be happening in this passage. First, Descartes reminds Mesland that on Descartes's own notion of indifference, it applies just to exercises of the lowest degree of freedom, where no

¹⁵ As Anthony Kenny points out, there is some doubt as to the date and the addressee of this letter (“Descartes on the Will,” pp. 25–26). But so far as I know, no one doubts that the author of the letter is Descartes.

compelling reasons drive us to one choice rather than another. Second, he recognizes that other people mean something very different by indifference, namely, simply the two-way power to chose between opposites. Third—and this is where the trouble begins—he asserts that “absolutely speaking,” though not “morally speaking,” we have this two-way power even when a very evident reason (presumably, a clear and distinct perception) moves us toward one of two opposite choices. Fourth—and this confirms the trouble—he asserts that we can always refuse to go along with a clear and distinct perception, in order to demonstrate our freedom of will. The last two points seem flatly incompatible with saying that we cannot refrain from assenting to clear and distinct perceptions at times when have them.

Let us take up these two points in reverse order. Of course, Descartes’s claim that we can “always” hold back from assenting to a clearly and distinctly perceived truth *can* be taken to mean that we can do so even at the very moment, or during the very span of time, that we perceive it clearly and distinctly. It would be silly to deny that Descartes’s words lend themselves to this interpretation. But the question is whether they must, or should, be read that way; for there is certainly an alternative reading, namely, that Descartes means only that we can always turn our careful attention away from a truth that we perceive clearly and distinctly, and that *then* we are no longer impelled to assent to it.

Now, there are at least three excellent reasons for adopting this alternative reading. First, it no longer saddles Descartes with contradicting his clearly stated position in the *Meditations*, *Principles of Philosophy*, *Objections and Replies*, *Correspondence* and elsewhere that clear and distinct perceptions are assent-compelling. After all, Descartes was if anything a highly subtle thinker, and in answering his critics he was also quite ready on occasion to appeal to straightforward common-sense distinctions, such as the distinction between denying or withholding *p* at the very same time that one clearly and distinctly perceives *p*, and denying or withholding *p* at times when one’s attention is not focused on *p*. For example, he appeals to such a common-sense distinction, albeit a different one in a different context, when he says to Gassendi: “when we say that an idea is innate in us, we do not mean that it is always there before us. This would mean that no idea was innate. We simply mean that we have within ourselves the faculty of summoning up the idea” (CSM II 132, SPW 131, AT VII 189). Likewise, then, Descartes may be interpreted as saying to Mesland that when we say that we are free to deny a clearly and

distinctly perceived proposition, we do not mean that we can deny it even when we are clearly and distinctly perceiving it, for this would mean that no proposition was assent-compelling. Rather, we simply mean that we have the power to turn our careful attention away from it, and that then we can deny it. That is both coherent and commonsensical. Why then should one insist on attributing to Descartes a contradiction, when there is an equally natural and perfectly straightforward way of reading his words that does not run into contradiction?

Second, Descartes's idea that we can "demonstrate the freedom of our will" by holding back from pursuing a known good or assenting to a clearly distinctly perceived truth (in our proposed sense of turning our careful attention away from these) makes good sense, for dissenting from a proposition that had only *some* evidence in its favor, or eschewing pursuit of something that had only limited value, would not be as clear a manifestation of our freedom of will as holding back from pursuing a good or assenting to a proposition that we remembered having recently very clearly and distinctly perceived. Indeed, by Descartes's lights this would appear to be the strongest possible demonstration of our freedom! By contrast, the notion that we can best demonstrate our freedom of will by withholding or denying a proposition at the very same time as we are perceiving it clearly and distinctly seems absurdly over-demanding.

Third, as Naaman-Zauderer acutely says,

Descartes could not have abandoned his view that clear and distinct perceptions compel assent without undertaking a fundamental revision of his most basic epistemological doctrines. One is his conception of error . . . Had Descartes believed that we can dissent from a clearly and distinctly perceived truth, he could not have argued that "it is quite impossible" for us to go wrong so long as we restrict the operations of the will to the clear and distinct perceptions of the intellect.¹⁶

This is of course correct, since Descartes would then have to admit, contra his claim that we cannot go wrong so long as we restrict our will to clear and distinct perceptions, that one possible way for us to go wrong is by denying a true proposition even though we clearly and distinctly perceive it. If we can deny a clearly and distinctly perceived proposition,

¹⁶ Naaman-Zauderer, Descartes' *Deontological Turn*, p. 122.

then we can deny a *true* clearly and distinctly perceived proposition, and so we can go wrong even with regard to something that we clearly and distinctly perceive. Our only complaint against Namaan-Zauderer, then, is not that she insists on the assent-compellingness of clear and distinct perceptions: this insistence is surely correct. Rather, it is that, like most commentators, she thinks that this assent-compellingness is incompatible with the principle of alternative possibilities, so that she is then in the position of having to maintain that assenting to a clear and distinct proposition can be free even though it violates that principle. By contrast, on our interpretation of the principle (or of how Descartes understands the principle), Descartes can maintain both the assent-compellingness of clear and distinct perception and the principle of alternative possibilities.

Let us turn to Descartes's remark that "when a very evident reason moves us in one direction, although morally speaking we can hardly move in the contrary direction, absolutely speaking we can." Here Descartes seems to be saying that in one sense ("morally speaking") we cannot deny a clear and distinct perception but in another sense ("absolutely speaking") we can. He does not explain what he means by "morally speaking" and "absolutely speaking." As Zaaman-Zauderer points out, his distinction cannot be understood in terms of the distinction he makes elsewhere between moral and absolute or metaphysical certainty. For moral certainty pertains to cases where there are more reasons on one side than on the other, whereas in a clear and distinct perception, all the reasons are on one side only.¹⁷ So, cases of moral certainty are not analogous to cases of being "morally speaking" unable to deny a clear and distinct perception, because in cases of moral certainty we are able to go against the preponderance of reasons and elect the weaker alternative, whereas in cases of being "morally speaking" unable to deny a clear and distinct perception, the reasons are conclusive and we are unable to go against them; thus it remains unclear what Descartes means when he says that "absolutely speaking," we can deny a clear and distinct perception.

Namaan-Zauderer considers but rejects two interpretations of what Descartes may have meant, by Lilli Alanen and Alan Nelson. According to Alanen, "when Descartes states that moving in the contrary direction would be morally impossible for us, he means only that it would be

¹⁷ Naaman-Zauderer, Descartes' *Deontological Turn*, p. 128 n49.

nonrational or morally unjustified.”¹⁸ Against this Naaman-Zauderer objects, reasonably enough, that

[For Descartes], rationality hinges on the (negatively formulated) duty to avoid assenting to matters we apprehend confusedly or obscurely. Unsurprisingly, therefore, Descartes does not explicitly prescribe consistent assent to our clear and distinct perceptions while attending to them. Such a prescription is not required because of the brute fact that psychologically or phenomenologically we cannot possibly do otherwise. So long as our attention is focused on clearly and distinctly perceived reasons, we can be wholly assured that we shall always assent to them.¹⁹

In other words, Descartes never gives voice to the idea that we morally *ought* to assent to clear and distinct perceptions, and this fact should not surprise us since he so explicitly says that current clear and distinct perceptions are irresistible, so that we have no choice but to assent to them. According to Nelson, on the other hand, “when Descartes writes that absolutely speaking we can hold back from admitting a clearly perceived truth, he is simply restating that the will can be diverted from a clear and distinct perception.”²⁰ (Presumably, then, when Descartes says that morally speaking we cannot hold back from admitting a clearly perceived truth, he simply means that we cannot do so while perceiving it clearly and distinctly.) Against this, Naaman-Zauderer objects that

Nelson’s interpretation does not fit well with Descartes’ suggestion that it is with regard to the same situation of choice that, “morally speaking,” we cannot act or judge against the weight of the evidence, whereas “absolutely speaking,” we can. This point was made by Ragland: “the ‘two senses’ passage maintains that, with respect to one and the same act of will, and at the same time, we can be both morally *unable* to hold back and absolutely *able* to hold back.”²¹

¹⁸ Naaman-Zauderer, Descartes’ *Deontological Turn*, p. 128.

¹⁹ Naaman-Zauderer, Descartes’ *Deontological Turn*, p. 129.

²⁰ Naaman-Zauderer, Descartes’ *Deontological Turn*, p. 128; quoting directly from Alan Nelson, “Descartes’ Ontology of Thought,” p. 172.

²¹ Naaman-Zauderer, Descartes’ *Deontological Turn*, p. 128. The quote from Ragland is from his “Descartes on the Principle of Alternative Possibilities,” p. 392.

Having rejected both Alanen's and Nelson's interpretations of Descartes's "morally speaking" versus "absolutely speaking" distinction, Naaman-Zauderer then goes on to propose her own interpretation:

In my understanding, moral necessity or impossibility, in this context, denotes a kind of practical or psychological impossibility, whereas absolute possibility signifies what is possible from an absolute, metaphysical point of view, irrespective of our inner experience. [Here Zaaman-Zauderer adds this footnote: "Descartes uses the phrase 'absolutely speaking' (*absolute loquendo*) or equivalents on other occasions as well, usually to signify the consideration of things from God's absolute point of view, as it were, independently of any actual limitations of our minds." See, e.g., letter to Clerselier, April 23, 1649 (AT V 355–54; CSMK 377); Second Replies (AT VII 145; CSM II 103)]. In saying that choosing otherwise when a very evident reason moves us in one direction is morally impossible for us, therefore, Descartes may mean that for us to do so is psychologically or practically impossible. From an absolute, purely metaphysical perspective, however, it is always possible for our will to act independently of our intellect, even in the face of clear and distinct ideas. The reason is that, metaphysically speaking, our intellect and will remain distinct from one another even when the intellect's perceptions are clear and distinct. Although we experience our intellect and our will as unified at the very moment of illumination—an experience that makes us morally (psychologically) incapable of acting against the intellect's overwhelming reasons—metaphysically or absolutely speaking we are capable of doing so.²²

This characterization of what is possible "absolutely speaking" is less than convincing, for when Descartes says to Mesland that "it is always open to us to hold back from pursuing a clearly known good, or from admitting a clearly perceived truth, provided we consider it a good thing to demonstrate the freedom of our will by so doing," he is surely talking about a real experiential possibility, not merely about an abstract one that can never be realized in experience, and that has to do only with the metaphysical distinction between the will and the intellect. How could an

²² Naaman-Zauderer, Descartes' *Deontological Turn*, p. 129–30.

act demonstrate to us our (real) freedom of the will, if that act occurred “irrespective of our inner experience” and went beyond “any actual limitation of our minds”? By contrast, our interpretation, according to which Descartes is claiming that we can demonstrate our freedom to ourselves by withholding even a truth that we remember having recently clearly and distinctly perceived, gives a perfectly literal, concrete content to Descartes’s claim. Naaman-Zauderer would presumably object, as she does against Nelson, that our interpretation does not do justice to the claim that it is the very same act of will that we are supposed to be morally unable to perform but absolutely able to perform. It seems, however, that this objection is not really coherent. For given that an actual, single act of the human will cannot intelligibly be said to belong both to the realm of that human being’s experience and to some abstract “metaphysical” realm that exists “irrespective of our inner experience” and “independently of any actual limitations of our minds,” it seems that the scenario she tries to envisage could only be one in which there is a unitary act (of holding back from assenting to a clear and distinct perception) that one is both able and unable to perform—and that is flatly contradictory.²³

Clearly, Descartes does mean to say at least that what is possible “absolutely speaking” is not always possible “morally speaking”—that the domain of what is possible absolutely speaking is broader than (though it includes) the domain of what is possible morally speaking. Our interpretation also fits well with this idea. For the class of propositions that can be clearly and distinctly perceived is broader than, and also includes, the class of propositions that are clearly and distinctly perceived by one finite mind at a particular time. So when Descartes says that absolutely speaking we can hold back from admitting a clearly perceived truth, “provided we consider it a good thing to demonstrate the freedom of our will by so doing,” he can plausibly be taken to mean that although we can hold back our assent from any proposition that falls within the “metaphysically” absolute totality of propositions that we could ever clearly and distinctly perceive (even from ones that we recently clearly and distinctly perceived), and may do so especially in order to demonstrate our freedom of will in the strongest way possible for us, nonetheless we cannot at any given time do this for the members of that class that also fall within the frequently altering sub-class of it that comprises

²³ Here I put to one side Immanuel Kant’s theory that the same action can belong both to the empirical world and to a noumenal world outside of space and time.

only the propositions that we are perceiving clearly and distinctly at that time.

There remains one question. In his *Replies to the Second Set of Objections*, as we saw in the previous chapter, Descartes says that some clear and distinct propositions are so simple that one cannot even think of them without clearly and distinctly perceiving them, so that, given the assent-compellingness of clear and distinct perceptions, one can never doubt these propositions at all:

Now some of these perceptions are so transparently clear and at the same time so simple that we cannot ever think of them without believing them to be true. The fact that I exist so long as I am thinking, or that what is done cannot be undone, are examples of truths in respect of which we manifestly possess this kind of certainty. For we cannot doubt them unless we think of them; but we cannot think of them without at the same time believing that they are true . . . Hence we cannot doubt them without at the same time believing that they are true; that is, we can never doubt them. (CSM II 104, M 105, SPW 141–2, AT VII 145–6)

How is this to be squared with our proposal that when Descartes says to Mesland that we can always hold back from assenting to a clear and distinctly perceived proposition, he means that we can do so provided we turn our careful attention away from it? It seems that with regard to these propositions, the only way that we can hold back our assent from them is by simply not thinking about them. We suggest that this consequence should be accepted, even though it may initially seem odd to say that we are “holding back” our assent from a proposition simply by not thinking about it. It might seem that to hold back our assent from a proposition *p*, we must be thinking about *p*, since to doubt that *p*, we must be thinking about *p*. But notice that holding back our assent from a clearly and distinctly perceived proposition is not the same as doubting it. One can hold back one’s assent from it by ignoring it, or by resolving not to think about any proposition that one can clearly and distinctly perceive, perhaps by filling one’s mind for five minutes with thoughts about only economics! This seems quite sufficient, strictly speaking, for it to be true that one is currently holding back from assenting to any clear and distinct proposition, including the *cogito* and the principle of noncontradiction, and that is all that Descartes’s position requires in order to be coherent.

6. Error and Evil

People generally regard human error as something bad or evil. We try to avoid error, and we highly value its opposite, knowledge. We erect systems of education, and a large part of the process of education consists in learning methods for avoiding error. In his *Fourth Meditation*, Descartes seeks not only to diagnose the cause of error and to prescribe the method for avoiding it, but also to show that the existence of this particular brand of evil is compatible with the existence of a perfect God. In that sense, the *Fourth Meditation* is Descartes's main contribution to the philosophical literature on the problem of evil. To put this contribution into a larger context, we shall begin with a general account of the problem of evil; subsequently, we shall explain how Descartes's views fit into this account.

6.1 The Problem of Evil

The classical problem of evil arises from the apparent contradiction between the existence of evil and the belief in God. The problem is that the existence of evil appears to be incompatible with the exalted attributes—omnipotence, omniscience, and supreme goodness—of the God of Judeo-Christianity. Does evil exist because God is unable to prevent it? Then he is not omnipotent. Does it exist because he does not want to prevent it? Then he is not supremely good. Does it exist because he is unaware of it? Then he is not omniscient. Thus, it seems that while God could have any two of the attributes that are traditionally taken as necessary to his perfection, he cannot have all of them. The problem can be formulated in the following logically valid “argument from evil”:

- (1) If God is omnipotent, then he can prevent evil.
- (2) If God is perfectly good, then he wants to prevent evil.
- (3) If God is omniscient, then he knows of any evil that exists.
- (4) Evil exists.
- (5) If evil exists, then either God is unable to prevent it, or he does not want to prevent it, or he is unaware of it.

.: (6) God is either not omnipotent, or not perfectly good, or not omniscient.

Traditionally, premises (3) and (5) are omitted because they are regarded as obvious and uncontroversial, and the last disjunct ("or" clause) of the conclusion is left understood as well. Since the argument is valid, it is not possible for all its premises to be true and its conclusion to be false. So a theist (a believer in God) must, on pain of contradiction, either accept the conclusion or reject at least one of the premises.

Could a theist simply accept the conclusion, that is, simply admit that the fact of evil proves that God is limited in power or in goodness or in knowledge? From a purely logical point of view, that is one option. Indeed, there are religions in which God is conceived as limited or finite in power and/or goodness. There are also religions, like Zoroastrianism, according to which there exist both a good deity and an evil one, who are each limited by the other's power, and who are in perpetual strife. At one point in David Hume's *Dialogues Concerning Natural Religion*, the character Cleanthes opts for a finite God, to which the character Philo, who is generally taken to represent Hume's own position, responds that even from a God whose power, goodness and wisdom were limited, an unbiased judge would expect a much better world than the actual world.²⁴ Be that as it may, the key point to recognize is that to accept the argument's conclusion would be to give up the belief in God as he is conceived in Judeo-Christianity. St. Anselm defined God in the twelfth century as "the being, than which none greater can be conceived," and Descartes defines him as a supremely perfect being. These definitions are not arbitrary: they answer to the concept of the God worshipped by Christians and Jews—a perfect God whose greatness is unsurpassable, and who therefore cannot be even a little bit limited in power, knowledge, or goodness. So even though the argument from evil is not explicitly directed against the existence of God, but purports only to prove that God cannot have all the divine attributes, its conclusion is tantamount to showing that God (at least the God of Judeo-Christianity) does not exist. Indeed the argument could be turned into an attempted disproof of his existence, just by adding the premise

- (7) If God exists, then he is omnipotent, perfectly good, and omniscient.

Then it follows from (6) and (7) that

- (8) God does not exist.

²⁴ David Hume, *Dialogues Concerning Natural Religion*, pp. 65–66.

Since theists (henceforth meaning by that term, believers in the Judeo-Christian God) cannot accept the argument's conclusion, (6), they must deny at least one of its premises. Let us consider first the possibility of denying (4), the premise that evil exists. What could such a denial mean? It might mean that (a) the things we call evil, such as pain, suffering, cruelty and so forth, do not exist, or (b) that these things exist, but are not really evil. Now (a) can be immediately dismissed, since it simply flies in the face of facts. Besides, we would then have to admit that at least the illusion of pain, suffering, and cruelty exist. But isn't that already an evil? Perhaps the existence of the illusion of pain, suffering, and cruelty isn't as bad as real pain, suffering and cruelty, but surely the absence of both real and illusory pain, suffering and cruelty would be still better. As for option (b), it fares no better. For if the things we consider to be evil, such as pain, suffering and cruelty, do not count as evil, then what would? If the answer is that nothing would even count as evil, then the term "evil" signifies nothing—it is meaningless. But then so is its opposite, "good," in which case it is meaningless to say that God is good. The point here is not that, as people sometimes say, there cannot be good if there is no evil; there is no reason to accept that claim. Rather, the point is that if nothing would even *count* or *qualify* as being evil, then the contrast between good and evil vanishes at the conceptual level, so that it becomes meaningless to call anything either "good" or "evil."

Can theists then reject premise (1)—that "if God is omnipotent, then he can prevent evil?" It may seem that they cannot, because by definition an omnipotent being can do anything. In other words, it may seem that (1) can be deduced from two analytic truths, as follows:

- (i) If God is omnipotent, then he can do anything.
- (ii) If God can do anything, then he can prevent evil.

. . . (1) If God is omnipotent, then he can prevent evil.

However, matters are not quite so simple, for (i) would commit us to saying that God can perform acts that involve a logical contradiction, for example, construct a four-sided triangle, or cause the mercury in a thermometer to be both one inch and two inches way from the bottom at the same time. Now as we have seen, Descartes did hold that an omnipotent could make contradictions true, so he would not be able to answer the argument from evil by objecting to (i). But, quite aside from the fact that Descartes's doctrine of the creation of the eternal truths would threaten

his doctrine that the human will closely resembles God's will, it seems quite incoherent. Aquinas's view that "nothing that implies a contradiction falls under the scope of God's omnipotence" seems to be the coherent view of omnipotence. It is worth pointing out, therefore, that (1) cannot really be avoided by objecting to (i), for (i) can easily be qualified so as to accommodate Aquinas's view of omnipotence, as follows:

- (ia) If God is omnipotent, then he can do anything that it is logically possible for him to do.

Presumably it would have been logically possible for God to prevent evil, so the argument for (1) can be amended by also qualifying (ii) to read this way:

- (iia) If God can do anything that it is logically possible for him to do, then he can prevent evil.

From (ia) and (iia), (1) follows as before.

So far, we have argued that a theist is committed to accepting premises (1) and (4). Premises (3) and (5) appear to be necessarily true, which is presumably why they are usually left understood; so the argument is commonly given an abbreviated formulation, which we shall adopt in the rest of our discussion:

- (1) If God is omnipotent, then he can prevent evil.
- (2) If God is perfectly good, then he wants to prevent evil.
- (3) Evil exists.

. . . (4) Either God is not omnipotent or he is not perfectly good.

It appears, then, that the only way to avoid the argument's conclusion is to reject premise (2). In other words, the theist must show that just because God is perfectly good, it does not follow that he *wants* to prevent all evil. This is indeed the main way that theists have tried to cope with the problem of evil: they have tried to show that the existence of evil is not incompatible with God's perfect goodness. This endeavor is called "theodicy," and Descartes's *Fourth Meditation* is, in part, a contribution to theodicy.

The first step in developing a theodicy is to draw a distinction between two kinds of evil: moral evil and nonmoral (or “natural”) evil. Moral evil is evil that results from human beings’ (and, some would add, angels’) wickedness or wrongdoing. Nonmoral evil is evil that does not result from human wickedness, but from causes like disease, plagues, earthquakes, tsunamis, droughts, and so forth. The problem for the theodicist then divides into two questions: (1) Why does a perfectly good God who is omnipotent allow the existence of moral evil, and (2) Why does such a God allow the existence of nonmoral evil?

Theodicy’s traditional response to the first question, commonly called “the free-will defense,” starts from the point that God, in his goodness, created a host of creatures possessing free will. Being omnipotent, he could of course have instead created only automata who always behaved as he wished them to. But this would not have been as good on his part as creating people with free will because, as the philosopher of religion John Hick puts it, “only persons [i.e., beings endowed with free will] could, in any meaningful sense, become ‘children of God,’ capable of entering into a personal relationship with their creator by a free and uncomelled response to his love.”²⁵ However, the theodicist’s response continues, the freedom of creatures has a price: it entails that there may well be moral evil. For free creatures cannot be guaranteed to always act rightly; rather, just because they are free, they may at times perform hurtful, wicked, morally wrong actions.

A natural objection to this line of thought is that the freedom of creatures implies the possibility of moral evil, but not its actuality. Why did God, if he is perfectly good as well as omnipotent, not create people who always freely choose to do what is morally right? In response, the theodicist should concede that freedom implies only the possibility of moral evil, not its actuality. Just because a creature is free to perform morally wrong actions, it does not follow that it will perform any such actions; conceivably, it might refrain from performing any such actions throughout its entire life. The picture changes, however, when the anti-theodicist adds that a perfectly good God would have created people who always freely choose to do what is right. For this amounts to saying that God would have caused or determined his creatures to freely choose only right actions, which is contradictory. If God causes or determines us to choose certain actions, then those actions cannot be completely free. Rather, we are like puppets, or people acting under post-hypnotic suggestion.

²⁵ John Hick, *Philosophy of Religion*, p. 39.

The theodicist concludes that it is no reflection on God's perfect goodness that moral evil exists. For the only way that God could have prevented it from existing would be by depriving his creatures of free will (or at least of a large measure of free will, since they could never choose morally wrong actions, but only choose among morally right ones), which would have undercut the possibility of moral goodness on their part and of their entering into a meaningful relationship with him. To put it differently, it is true that our freedom implies only the possibility of moral evil, not its actuality. But our freedom also implies that God is not responsible for any moral evil that actually occurs, for to hold him responsible for it would be to say that he should have prevented it. But if he had prevented it, he would also have deprived his creatures of their freedom, or significantly reduced it.

The free-will defense, even if successful, provides only a partial solution to the problem of evil, for it leaves untouched the problem presented by nonmoral evil; that is, by physical suffering and mental anguish that do not result from human wickedness or wrongdoing. At best, then, the free-will defense forces a religious skeptic to reformulate the argument from evil only in terms of nonmoral evil, as follows:

- (1) If God is omnipotent, then he can prevent nonmoral evil.
 - (2) If God is perfectly good, then he wants to prevent nonmoral evil.
 - (3) Nonmoral evil exists.
-
- .: (4) Either God is not omnipotent or he is not perfectly good.

There are two different traditional responses to this argument. The first, which stems from St. Augustine (345–430 A.D.) is directed against premise (3); the second, which stems from Irenaeus (130–202 A.D.), is directed against premise (2). We shall consider them in turn.²⁶

The Augustinian theodicy holds, basically, that there is no such thing as nonmoral evil; the only kind of evil that exists is moral evil, which is a kind of turning away from God. Augustine's denial of nonmoral evil rests

²⁶ The discussion that follows in this section is strongly influenced by the writings of John Hick, cited in notes 25 and 28. For a penetrating critique of Hick's Irenean theodicy, see Edward H. Madden and Peter H. Hare, *Evil and the Concept of God*, especially pp. 83–90, 85–84, 102–3.

on two pillars: his privative view of evil, and his aesthetic view of evil. The privative view is that nonmoral evil is not something real or positive; instead it is merely an absence or lack of perfection. According to Augustine, everything that exists is good, but some things are better or more nearly perfect than others. Most obviously, God, the creator, is perfect, but created things are less than perfect; if they were perfect, they would not be any different from God. Thus it is in the very nature of created things to be imperfect, but this does not mean that anything that God created is positively evil—all things are good. This view, which comes from Plotinus and ultimately from Plato, equates being with goodness. It implies that just as there are degrees of goodness, there are degrees of being, and that just as nonbeing is unreal, likewise nonmoral evil is unreal.

The aesthetic view of nonmoral evil, which complements the privative view and also derives from Plotinus, is that the things that we call evil only appear to us as evil because we do not understand them in relation to the whole creation. If we were to see them in relation to the whole, then we would appreciate the fact that every part of creation, no matter how it appears from our limited human perspective, contributes to the goodness of the whole. As an analogy, consider a single measure of a Bach fugue, heard in isolation from the rest of the composition. It can be dissonant and hence “ugly,” but it may be an indispensable element for enhancing or maximizing the beauty of the whole piece.²⁷ Likewise, a single patch of color on a painting might utterly lack aesthetic appeal, but significantly enhance the beauty of the whole painting.

We shall say more about the Augustinian theodicy when we return to Descartes, but first let us turn our attention to the Irenean theodicy, especially as defended by its modern advocate, John Hick. Irenean theodicy does not deny the existence of nonmoral evil. Rather, it opposes premise (2) of the above argument—the premise that if God is perfectly good, then he wants to prevent nonmoral evil—on the ground that the nonmoral evil that exists is required for the purpose of bringing about a greater good than would be possible without it. Specifically, there are certain valuable human virtues that could not be developed or manifested in a world lacking all nonmoral evil, such as courage, perseverance, sympathy, charity, protectiveness, and so forth. In a world devoid of any adversity or obstacles, there could never be any opportunity to exercise any of these virtues.

²⁷ Wallace I. Matson, *The Existence of God*, p. 146.

This line of thought can be supplemented by attacking what seems to be an assumption behind premise (2). This is the assumption that a perfectly good God's purpose in creating the world must have been to produce a pleasure paradise for humans to inhabit, akin to building a cage for some favorite pet animal. In other words, his aim, since he is perfectly good, must have been to provide humans with as safe, pleasurable, and nondemanding an environment as possible. Irenean theodicy's response to this assumption is that it misrepresents God's purpose. His true purpose was to create a world of "soul-making," in which creatures would develop and exercise the nobler virtues, and for this purpose it is essential that there be hardships, obstacles, and adversity to overcome. The nobler virtues could not get a foothold if nonmoral evils did not exist.

How might Irenean theodicy, as so far described, be criticized? Well, a critic can point out that although it is true that *sometimes* nonmoral evil breeds virtues like courage, patience, compassion, steadfastness, and so on, sometimes it does just the opposite. Some people crumble under the weight of adversity; some become permanently embittered, or even go insane. Thus, a critic could reformulate the argument from evil this way:

- (1) If God is omnipotent, then he can prevent nonmoral evil that does not help to produce a greater good.
- (2) If God is perfectly good, then he wants to prevent nonmoral evil that does not help to produce a greater good.
- (3) Nonmoral evil that does not help to produce a greater good exists.

. . . (4) Either God is not omnipotent or he is not perfectly good.

In his book *Evil and the God of Love*, John Hick attempts to refute even this carefully qualified version of the argument. He does so by appealing to one of the most characteristic beliefs of Christianity, namely the belief in life after death.²⁸ Hick argues that although the evils that we encounter during life on earth do not all contribute to a greater good, they may well do so if we take into consideration the afterlife. He holds that the belief in an afterlife in which the evils suffered in this world ultimately lead to a greater good should be an intrinsic part of

²⁸ John Hick, *Evil and the God of Love*, pp. 373–77.

the Irenean theodicy that he favors. Notice that for the purpose of answering the argument from evil, it is not necessary to affirm that there will be an afterlife. All that is necessary, strictly speaking, is to hold that there *may* be an afterlife or, more accurately put, that it is not unreasonable or irrational to affirm this. For this maneuver forces the argument from evil's proponent to qualify the argument once again, so that it goes as follows:

- (1) If God is omnipotent, then he can prevent nonmoral evil that does not help to produce a greater good in either this life or the afterlife.
 - (2) If God is perfectly good, then he wants to prevent nonmoral evil that does not help to produce a greater good in either this life or the afterlife.
 - (3) Nonmoral evil that does not help to produce a greater good in either this life or the afterlife exists.
-
- . . . (4) Either God is not omnipotent or he is not perfectly good.

However, at this point the theodicist could claim that the argument form evil has become unreliable, because nobody has adequate evidence for premise (3); nobody can affirm with justifiable confidence that there is not an afterlife during which whatever nonmoral evil we endured in our earthly life will lead to a greater good. Thus, the theodicist's best hope may lie in forcing the religious skeptic to qualify premise (2) of the argument from evil to the point where premise (3) becomes so dubious that the argument becomes unreliable.

To be sure, there are possible responses that a skeptic can still make. One is that some nonmoral evils suffered in this life are so crushing that the mere memory of them would ruin or at least damage the quality of the after-life. Another is to point to the fact of animal suffering: non-human animals, too, are sentient, conscious beings, and multitudes of them suffer great pain and distress. Why would a perfectly good God allow this? Are we to believe that animals, too, have an afterlife in which their past suffering leads to a greater good? Yet another response would be to argue that the belief in an afterlife is unreasonable. After all, we know that when certain parts of the brain are damaged, corresponding mental functions and forms of consciousness cease to occur. Is it not highly probable, then, that when brain activity stops altogether, so does

all mental functioning and consciousness? We leave further reflection on these questions to the reader.

6.2 Cartesian Theodicy

Against the background the general problem of evil, let us turn to Descartes's attempt to show that the existence of error is compatible with the existence of an omnipotent, perfectly good, omniscient God. First, it is obvious that Descartes treats human error in the same way as traditional theodicy treats moral evil, for he applies the free-will defense to it. Human error results from a certain misuse of our free will, namely that of making judgments regarding matters that we do not perceive clearly and distinctly, and it can be avoided by restricting our judgments to matters that we do perceive clearly and distinctly. Moreover, Descartes treats erroneous judgments in tandem with morally bad choices, and his language implies that he regards mistakes about what is true or false as morally blameworthy, as when he says that even if through sheer luck I make a true judgment about something that I do not perceive clearly and distinctly, "I shall still be at fault since it is clear by the natural light that the perception of the intellect should always precede the determination of the will" (CSM II 60, SPW 103, AT VII 60).

Descartes also appeals to the privative view of evil: not without precedent, he extends the privative view beyond nonmoral evil to (what he regards as a species of) moral evil. Thus near the beginning of the *Fourth Meditation*, he says

I realize that I am, as it were, something intermediate between God and nothingness, or between supreme being and non-being: my nature is such that in so far as I was created by the supreme being, there is nothing in me to enable me to go wrong or lead me astray; but in so far as I am not myself the supreme being and am lacking in countless respects, it is no wonder that I make mistakes. I understand, then, that error as such is not something real which depends on God, but merely a defect. Hence my going wrong does not require me to have a faculty specially bestowed on me by God; it simply happens as a result of the fact that the faculty of true judgment which I have from God is in my case not infinite. (CSM II 38, SPW 99–100, AT VII 54)

Here Descartes is preparing the way for his view that neither of the faculties used in making judgments, the intellect and the will, have error

“built in” to them, so that they are in no way defective. But the reason he gives for this, that “error is not something real,” expresses the privative view of evil.

In the next paragraph, however, Descartes finds that this appeal is insufficient:

But this is not entirely satisfactory. For error is not a pure negation, but rather a privation of lack of some knowledge which somehow should be in me. (CSM II 38, SPW 100, AT VII 55)

Here Descartes makes a distinction between a mere negation or lack, and a privation. A mere lack is the absence of a desirable characteristic that a thing does not have because of its very nature; a privation is the absence of a desirable characteristic that thing could have had while keeping the same nature, and in that sense “ought” to have. For example, the absence of eyes in certain deep-sea fish that live in total darkness is a lack but not a privation, but the absence of functioning eyes in a human being would be a privation. We say that a sightless deep-water fish lacks sight but not that it is *deprived* of sight, but we say of a human being who lacks functioning eyes that she is deprived of sight and suffers from blindness. Now, Descartes asks, why isn’t the occurrence of human errors a privation rather than a mere lack? His answer is that, apart from the fact that we cannot fathom God’s every purpose,

Whenever we are inquiring whether the works God are perfect, we ought to look at the whole universe, not just one created thing on its own. For what would perhaps rightly appear very imperfect if it existed on its own is quite perfect when its function as part of the universe is considered. (CSM II 39, SPW 100, AT VII 55)

This is essentially an appeal to the aesthetic conception of evil: parts of a whole that may have little or no value in isolation may well contribute significantly to the goodness of the whole. Thus, in his attempt to reconcile the existence of human error with God’s perfection, Descartes appeals not only to the free-will defense, but also to both strands of the Augustinian theodicy, the privative view of evil and the aesthetic view of evil.

Even if these appeals can solve the theodicean problem caused by the actual occurrence of human errors, there remains a question: why did a

perfect God create humans who are even *prone* to error? Why did he not, instead, give us a nature such that we never go wrong? As Descartes puts it:

There is . . . no doubt that God could have given me a nature such that I was never mistaken [since he is omnipotent]; again, there is no doubt that he always wills what is best [since he is perfectly good]. Is it then better that I should make mistakes than that I should not do so? (CSM II 38, SPW 100, AT VII 55)

God could easily have brought it about that without loosing my freedom, and despite the limitations in my knowledge, I should nonetheless never make a mistake. He could, for example, have endowed my intellect with a clear and distinct perception of everything about which I was ever likely to deliberate. . . . Had God made me this way . . . I would have been more perfect than I am now. (CSM II 42, SPW 104, AT VII 61)

As these passages show, Descartes readily admits that God could have made humans incapable of error simply by giving them an understanding with as great a scope as their wills. So, why did he not do so?

Desacartes's response to this question is a straightforward appeal to the aesthetic conception of evil:

I cannot . . . deny that there may in some way be more perfection in the universe as a whole because some of its parts are not immune from error, while others are immune, than there would be if all the parts were exactly alike. (CSM II 43–44. SPW 104, AT VII 61)

In other words, in the big picture—from the point of view of the whole universe—it may be better if some beings are prone to error than if all beings were infallible. In his Reply to Gassendi's objections, Descartes explains this point with a striking analogy:

A . . . comparison [could be made] . . . between someone who wanted the whole of the human body to be covered with eyes so as to look more beautiful (there being no part of the body more beautiful than the eye), and someone who thinks that there ought not to have been any creatures in the world who were liable to error (i.e., not wholly perfect). (CSM II 258–9)

It seems, then, that Descartes's main way of responding to the question of why a perfect God did not make humans immune to error is to appeal to the aesthetic conception of evil.

Is there any place in Descartes's thinking about error for the Irenean theodicy? It seems not. The idea that we could perfect our souls by learning from our errors, or by gradually increasing our knowledge and wisdom through overcoming obstacles to the acquisition of knowledge, seems foreign to Descartes's thinking, despite his own life-long, prodigious efforts to increase his knowledge. Thus, in the *Third Meditation* he considers the possibility that he might be a greater being than he seems to be, and even so great as to be able to cause the idea of God, because his knowledge increases gradually and there is no reason why it could not grow indefinitely. His response is that the very fact that his knowledge increases gradually shows how imperfect he is: "this gradual increase in knowledge is itself the surest sign of imperfection" (CSM II 32, SPW 95, AT VII 47). This seems incompatible with the idea that a soul which has had to develop its intellectual virtues "the hard way" is more nearly perfect than one to whom everything "comes easy."

6.3 Some Critical Reflections

In this section, we shall offer some brief critical reflections on the theodicy that underlies Descartes's attempt to reconcile the fact that humans are prone to error with God's perfect goodness. Since Descartes does not appeal to the Irenean theodicy, we will focus only on the Augustinian themes that we identified in Cartesian theodicy—the privative and the aesthetic conceptions of evil.

Few philosophers nowadays would look with favor on the equation of being and goodness that lies at the heart of the privative conception of evil. Contemporary philosophers commonly make a distinction, which they regard as fundamental, between "facts" and "values." This fact/value distinction means that for anything that exists, it is an open question whether it is good, bad, or indifferent—a question that can be reasonably answered only on the basis of some theory of value or some ethical or aesthetic principles. Thus, the distinction contradicts the idea that anything is good, or valuable, just because it exists, or "has being." But as foreign as the equation of being and goodness may be to much contemporary philosophy, it is at the heart of Judeo-Christianity. Christians and Jews believe that the world was created by and is looked over by a perfectly

good God. From their point of view, such a world must be good, and to discover that it was not good would be devastating and terrifying. It would threaten the very "meaning" of their lives, since they could no longer see humans as children of a good deity who cares for them and will "deliver us from evil" and from eternal oblivion. Thus, the existence of evil cannot be seen by Christians and Jews as a trivial blemish or a small imperfection; instead, it confronts them with a formidable problem. No wonder, then, that one classic response to the problem is deny the reality of evil, and to see the belief that evil is real as based on some sort of misunderstanding of God's purposes.

All the same, however, the existence of horrible and prolonged physical pain, debilitating illness, extreme mental suffering and anguish, and of the natural forces and catastrophes that cause these, is obvious. Epidemics, droughts, earthquakes, tsunamis, tornadoes, agonizing and fatal cancers, horrible work-related injuries, degenerative physical and mental illnesses are just a few examples. The existence of these things presents religious believers with a wrenching dilemma: either admit that they are genuinely evil and utterly inexplicable by us as being works of God, or deny that they are really evil. If one opts for the first horn of the dilemma, then one's religious belief seems to ignore the dictates of reason. If one opts for the second horn, then one seems wholly callous: how can one really believe, for example, that the painful death of a child from inoperable throat cancer is not evil? Further, even the notion that there is moral evil then seems imperiled, since if pain and suffering are not really evil in themselves, it is hard to see how intentionally causing them should be evil. Are we to say that they are evil when intentionally caused by humans, but not when intentionally caused by God?

The aesthetic conception of evil is perhaps best seen as an attempt to shore up the privative conception of evil, by denying the ultimate reality of evil and thereby softening the second horn of the above dilemma. To answer the question of how an innocent child's suffering and death can be seen as only a privation of good and not really an evil, the aesthetic conception of evil would suggest that the child's suffering and death be seen as merely an absence of good in one part of the universe that is compensated for by the resulting goodness of the whole.

This suggestion, however, overlooks an important point. The aesthetic conception of evil turns on an analogy between parts of an aesthetically pleasing whole like a work of art and conscious beings who are parts of God's whole creation. But the analogy is flawed, for a part of an

aesthetically pleasing whole, like a single measure of a Bach fugue or a portion of a painting, is not a sentient being capable of suffering or joy. So, in judging the goodness of the fugue or the painting as a whole, the only point of view that is relevant is that of the listener or the viewer. But in judging the goodness of the whole creation, we must also consider the points of view of each of its living, conscious parts.²⁹

In light of the fragility of the aesthetic conception of evil, it may now seem that Descartes's attempt to "explain away" the fallibility of humans rests on very weak ground. But perhaps we can offer an alternative response that Descartes does not give but that he could have given. He could have held that God gave us an understanding as wide as our wills, so that we are indeed *capable* of attaining clear and distinct perceptions on all matters whatsoever. Then we would still have the freedom to err. But this freedom would not be based on the mismatch between our intellect and our will—the will would not extend beyond the intellect. Rather, our freedom to err would stem only from our freedom not to engage in the arduous work that would still be necessary for finite creatures to attain clear and distinct perceptions regarding all matters. But availing ourselves of this freedom would then not be a nonmoral evil, but rather a moral evil manifesting a form of laziness or a lack of perseverance. Thus, the capacity to make errors, no less than the actual commission of errors, would reflect only a moral flaw in humans, and Descartes would not need the aesthetic (nor even the privative) conception of nonmoral evil to account for it. Of course adopting this alternative would not help to solve the problem posed by nonmoral evil, and it would require Descartes to drop the view that there are some matters that the human intellect, no matter how assiduously we use it, simply cannot grasp. We leave it to the reader to consider whether this would have been a viable option for him.

²⁹ Matson, *The Existence of God*, pp. 147–48.

This page intentionally left blank

Meditation V

The Ontological Argument for the Existence of God

1. Descartes's Ontological Argument

The Ontological Argument was originally put forward by St. Anselm (1033–1109), who was Archbishop of Canterbury under William II and one of the greatest Christian theologians. Anselm began by defining God as “a being than which nothing greater can be conceived”—a definition that beautifully expresses the Judeo-Christian concept of God as an absolutely unsurpassable being. He then asked whether it is possible that this being exists only in our minds, as a mere object of thought. He answered that this is impossible, for then this very being than which nothing greater can be conceived would be a being than which a greater *can* be conceived; for if it had also existed in reality, then *it* would have been greater. Anselm concluded that God exists both as an object of thought and in reality.

Although Descartes does not mention Anselm, and it is unclear whether he was familiar with the *Proslogium* (the little book in which Anselm advanced his argument), Descartes's argument for God's existence in *Meditation V* is essentially a modernized version of Anselm's argument. Instead of defining God as a being than which nothing greater can be conceived, Descartes defines God, more simply, as “a supremely perfect being.” He then argues that it is impossible for such a being not to exist, because then it would not be supremely perfect! Before we look at this arresting piece of reasoning more closely, notice how well it suits Descartes's purposes. As already mentioned, Descartes's arguments for God's existence must not employ any premises that refer to the material world, or indeed to anything existing outside his own thought, because all such things are still doubtful. The Ontological Argument satisfies this requirement perfectly, because it turns on the mere definition or concept of God as an absolutely unsurpassable being.

Descartes introduces the argument some way into *Meditation V*, as follows:

But if the mere fact that I can produce from my thought the idea of something entails that everything which I clearly and distinctly perceive to belong to that thing really does belong to it, is not this a possible basis for another argument to prove the existence of God? Certainly, the idea of God, or a supremely perfect being, is one which I find within me just as surely as the idea of any shape or number. And my understanding that it belongs to his nature that he always exists is no less clear and distinct than is the case when I prove of any shape or number that some property belongs to its nature. Hence, even if it turned out that not everything on which I have meditated in these past days is true, I ought still to regard the existence of God as having at least the same level of certainty as I have hitherto attributed to the truths of mathematics. (CSM II 45, SPW 107, AT VII 65–6)

This passage's opening sentence reveals that the *Meditation V* argument for God's existence, unlike the *Meditation III* arguments, does not purport to provide a vindication of the clarity-and-distinctness criterion of truth. Instead, the *Meditation V* argument explicitly uses that criterion. This difference in the roles of the arguments is also revealed by Descartes's saying that the *Meditation V* argument shows that God's existence has "at least the same level of certainty as I have hitherto attributed to the truths of mathematics." This statement indicates that Descartes does not see the *Meditation V* argument as somehow attaining a higher level of assurance than mathematics, which would then enable it to guarantee the truths of mathematics. Instead, he sees the argument as resting, just like mathematics, on the principle that clear and distinct perceptions are true.

The fact that Descartes uses his criterion of truth in *Meditation V* is not surprising; for by the time he reaches that *Meditation*, he takes himself to have already established the criterion. Thus, as Copleston has argued, the problem of the circle does not arise for the *Meditation V* argument, as it did for the *Meditation III* arguments.¹ It is true, as Copleston also points out, that in the *Principles of Philosophy* Descartes presents the Ontological Argument before the causal ones and derives his criterion of

¹ Frederick Copleston, *A History of Philosophy*, pp. 105–106.

truth afterward. So in that work all the theological arguments are evidently seen as contributing to the guarantee of clear and distinct perception. Perhaps, then, the difference in the way Descartes treats the arguments in the *Meditations* does not reflect a deep-seated feature of his thought. But the difference is there and should be noted.

Let us now turn directly to Descartes's Ontological Argument itself. The passage that we have quoted from *Meditation V* does not make the argument's structure very clear. But Descartes explained his reasoning more fully in his *Replies* to both the first and second sets of *Objections*:

My argument however was as follows: 'That which we clearly and distinctly understand to belong to the true an immutable nature, or essence, or form of something, can be truly asserted of that thing. But once we have made a sufficiently careful investigation of what God is, we clearly and distinctly understand that existence belongs to his true and immutable nature. Hence we can now truly assert of God that he does exist.' Here at least the conclusion does follow from the premises. But, what is more, the major premise cannot be denied, because it has already been conceded that whatever we clearly and distinctly understand is true. (CSM II 83, M 100, SPW 136–7, AT VII 115–6)

But my major premise was this: 'That which we clearly understand to belong to the nature of something can be truly affirmed of that thing.' Thus, if being an animal belongs to the nature of man, it can be affirmed that man is an animal; and if having three angles equal to two right angles belongs to the nature of a triangle, it can be affirmed that a triangle has three angles equal to two right angles; and if existence belongs to the nature of God, it can be affirmed that God exists, and so on. Now the minor premise of my argument was: 'yet it belongs to the nature of God that he exists.' And from these two premises the evident conclusion to be drawn is the one which I drew: 'Therefore it can truly be affirmed of God that he exists.' (CSM II 106–7, AT VII 149–150)

Drawing on these passages, we may formulate Descartes's argument as follows:

- (1) Whatever I clearly and distinctly perceive to belong to the nature or essence of a thing can be truly affirmed of that thing.

(2) I clearly and distinctly perceive that existence belongs to the nature or essence of a supremely perfect being.

.: Existence can be truly affirmed of a supremely perfect being, i.e., a supremely perfect being exists.

For brevity's sake, we have left understood the phrase "true and immutable" that Descartes puts in front of "nature or essence" in both premises. The point of this phrase will be seen in the next section.

Premise 1 is based on two ideas: Descartes's clarity-and-distinctness criterion of truth and the connection (emphasized in the last-quoted passage) between a property's belonging to the essence of a thing and belonging to (i.e., being such that it can be "truly asserted of" or "truly affirmed of") the thing itself. Thus, we can construct the following subargument for premise 1:

Whatever I clearly and distinctly perceive to belong to the nature or essence of a thing does belong to its nature or essence.

Whatever belongs to the nature or essence of a thing can be truly affirmed of that thing.

.: Whatever I clearly and distinctly perceive to belong to the nature or essence of a thing can be truly affirmed of that thing.

The first premise of this subargument is a direct application of Descartes's criterion of truth. The second premise is an analytic truth. For "essence" and "nature" here mean basically the same as "definition." But if a property *P* belongs to the very definition of a thing, then, of course, *P* can be truly affirmed of (i.e., must actually belong to) that thing itself.

Premise 2 of the main argument involves the key idea in the Ontological Argument—the connection between *supreme perfection* and *existence*. Descartes invokes this connection in replying to the first of three possible objections that he considers immediately after introducing the argument. The first of these possible objections may be paraphrased: "I have always made a distinction between the essence and the existence of a thing, between the question '*What* is *X*?' and the question '*Does X exist?*' So why can't I make this distinction in the case of God, and thereby conceive of God as not existing?" Descartes's reply is that on considering

the matter more carefully, it becomes obvious that the existence of God cannot be separated from his essence anymore than having its angles equal 180 degrees can be separated from the essence of a triangle or having a valley can be separated from the essence of a mountain. (Descartes clarified the second example in a letter of January 19, 1642 to Gibieuf explaining that by a mountain he had meant merely an uphill slope, and by a valley a downhill slope [CSMK 202, AT III 477].) Descartes is here saying that there is an obvious, logical connection between being God (having God's "essence" or "nature") and existing, just as there is an obvious, logical connection between being a triangle and having three angles that equal 180 degrees or having an uphill slope and having a downhill slope.

Why is there such a connection between the essence and the existence of God? Descartes's answer comes in his statements that

it is just as much of a contradiction to think of God (that is, a supremely perfect being) lacking existence (that is, lacking a perfection), as it is to think of a mountain without a valley. (CSM II 46, SPW 107, AT VII 66)

and

I am not free to think of God without existence (that is, a supremely perfect being without a supreme perfection). (CSM II 46, SPW 107, AT VII 67)

In effect, Descartes is here giving a subargument for the "that"-clause of premise (2) of the main argument, which is the part of the premise in which the crucial connection between supreme perfection and existence is asserted. This subargument is:

All perfections belong to the essence of a supremely perfect being.

Existence is a perfection.

∴ Existence belongs to the essence of a supremely perfect being.

This little argument expresses the key idea in Descartes's Ontological Argument—that the nature or essence of a supremely perfect being must include existence, *because such a being must have all perfections and existence is*

itself a perfection. In other words, since existence is a perfection, anything that failed to exist would not have *all* perfections and so could not be a supremely perfect being. This is why there is an obvious, logically necessary connection between being God (i.e., being a supremely perfect being) and existing.

In order to make his argument still clearer—to, as it were, “teach” it to us—Descartes considers two further possible, hypothetical objections to it. The first may be paraphrased:

Although it is true that I cannot think of God without existence anymore than of a mountain without a valley, it would be invalid to argue: “I think of a mountain with a valley; therefore, there exists a mountain with a valley.” Likewise, why isn’t it invalid to argue: “I think of God as existing; therefore, God exists”? After all, my thought imposes no necessity on anything: just because, for example, I attribute wings to a horse in my imagination, it does not follow that a winged horse exists; likewise, just because I attribute existence to God in my thinking, it does not follow that God really exists.

In Descartes’s reply to this objection, he grants that it would be invalid to argue:

I cannot think of a mountain without a valley.

. . . A mountain and/or a valley exists.

However, he claims, it is valid to argue:

I cannot think of a mountain without a valley.

. . . A mountain is inseparable from a valley.

(Of course, this assumes the general principle, which the objection does not call into question, that if I cannot think of [i.e., clearly and distinctly perceive] a thing X without a property P, then P is inseparable from X.) Therefore, Descartes continues, it is also valid to argue:

- (i) I cannot think of God without existence.
- (ii) Existence is inseparable from God. (from step i)
- (iii) God really exists. (from step ii)

Furthermore, he adds, it is the necessity of God's existence that determines my thought, not vice-versa; for while I can think of a horse either with or without wings, I cannot think of God (i.e., of a supremely perfect being) without existence (i.e., without this particular perfection).

The last possible objection Descartes considers is a little more complex. It goes as follows. If one supposes or assumes that

(1) A supremely perfect being has all perfections

then, since

(2) Existence is a perfection,

it follows that:

(3) A supremely perfect being exists.

However, the objection continues, the original supposition or assumption, (1), is no more necessary than the assumption that

(1') A circle can inscribe all quadrilateral figures

which, since

(2') A rhombus is a quadrilateral figure,

would lead to the false conclusion that

(3') A circle can inscribe a rhombus.

Since a rhombus is an oblique-angled parallelogram, it cannot be inscribed in a circle (i.e., placed within a circle with all four corners of the rhombus touching the circle's circumference). This is why (3') is false. The point of the objection is that just as the argument from (1') and (2') to (3') must be unsound, so the argument from (1) and (2) to (3) is unsound.

Descartes replies, as one would expect, by rejecting the analogy between the two arguments—specifically the analogy between (1) and (1'). We may paraphrase his reply:

It's true that I do not have to think about a supremely perfect being, anymore than about a circle. In that sense, and in that

sense alone, (1) and (1') are alike, are both “not necessary.” But in all other important respects, they are different. For whenever I do think of a supremely perfect being, I perceive clearly and distinctly that (1) is necessarily true (just as when I think of a triangle, I perceive clearly and distinctly that it must have three angles). But when I think of a circle, I do not, and indeed cannot, clearly and distinctly perceive that (1') is necessarily true. On the contrary, (1') is false, so no wonder that the falsehood (3') can be derived from it.

Although the objection just considered may have looked weak even before Descartes’s rebuttal, it is important. For in setting out this objection, Descartes has in effect given us a very simple version of the Ontological Argument itself. This simple version consists of just the first three numbered statements from the objection:

- (1) A supremely perfect being has all perfections.
- (2) Existence is a perfection.

∴ (3) A supremely perfect being exists.

This concise Ontological Argument, like the subargument for the “that”-clause of premise (2) of the longer argument (see page 225), explicitly uses the premise “Existence is a perfection.” This idea, which will be examined in the next section, is a crucial common element of Descartes’s and Anselm’s ontological arguments. For both Anselm and Descartes, existence contributes to a thing’s “greatness” (Anselm) or perfection (Descartes). So, having defined God, in accordance with Judaism and Christianity, as a supremely perfect being (Descartes), or as one than which nothing greater can be thought (Anselm), one is logically compelled to say that such a being exists. This is really the heart of the Ontological Argument.

2. Critique of the Ontological Argument

Most philosophers today agree in rejecting the Ontological Argument, but there is less agreement as to exactly what is wrong with it. In this section, we shall weigh three different objections to the argument.

2.1 Gaunilo's Objection

The first objection to the Ontological Argument was made by a monk named Gaunilo against Anselm's original argument. Gaunilo objected that if Anselm's argument were sound, then the same reasoning could be used to "prove" the existence of things that do not exist, such as an island than which none greater can be conceived. As applied to Descartes's argument, Gaunilo's objection might go as follows. Suppose that for premise (2) of Descartes's initial argument, we substitute the premise that:

(2') I clearly and distinctly perceive that existence belongs to the nature or essence of a most perfect island (most perfect lion, most perfect cigar, most perfect mustache, etc.).

Then from (2') and premise 1 (that whatever I perceive clearly and distinctly to belong to the nature or essence of a thing can be truly affirmed of that thing) we can deduce that there really exists a most perfect island, a most perfect lion, a most perfect cigar, a most perfect mustache—indeed a most perfect thing of any and every type! Surely, this absurd consequence shows that something is wrong with Descartes's argument.

To meet this objection, Descartes would probably have appealed to the theory about natures or essences that he sketches at the beginning of *Meditation V*, before stating the Ontological Argument itself. There Descartes says that he has ideas of certain things that whether or not they actually exist and whether or not he even thinks of them, have natures or essences of their own, which he has not invented and cannot change. Descartes calls such natures or essences "true and immutable natures." As an example, Descartes cites a triangle. He points out that this geometrical figure has a nature or essence which he did not invent and which does not depend in any way on his mind, since various properties can be proved of it, for example, that its longest side is opposite its widest angle, that its three angles equal two right angles, and so forth. For Descartes, this example is merely one illustration of a very general view—that geometrical figures and other mathematical objects, such as numbers, have true and immutable natures or essences, which account for the certainty of mathematics. Furthermore, at least one nonmathematical object, namely, the supremely perfect being, has a true and immutable nature or essence—one that differs from all other essences in that it alone includes existence. By contrast, Descartes believes, a most perfect island, lion or cigar has *no* true and immutable nature or essence.

Such things are merely “fictitious” creatures invented by ourselves, like the centaur (a mythical beast having the head, arms, and torso of a man but the body and legs of a horse). Consequently, while premise 2 of the Ontological Argument is true, (2') is false (inasmuch as “nature or essence” in both of these statements is short for “true and immutable nature or essence”). Therefore, the argument cannot be used to “prove” the existence of most perfect islands or lions or cigars.

This reply to Gaunilo’s objection, however, is not without difficulty; for it places an onus on Descartes to give a criterion for distinguishing between things that do have a “true and immutable nature or essence” and things that do not. Why does a triangle have a true and immutable nature, whereas (say) a centaur does not? Descartes would doubtless say that while the centaur is invented by us and dependent on our thought, the triangle is not; and at first this may sound good. But what does it really mean? It does not mean that while centaurs don’t really exist in nature, triangles do; for Descartes explicitly says that even if no triangle exists or has ever existed outside his thought, this geometrical figure still has a true and immutable nature. So there seems to be no relevant difference here; for just as the concept of a centaur exists only thanks to a definition constructed by humans that specifies the properties of a centaur, so the concept of a triangle exists only thanks to a definition constructed by humans that specifies the properties of a triangle. And just as, having once defined a triangle, we can demonstrate various properties of that figure (e.g., that its three angles equal 180 degrees), so, having once defined a centaur, we can demonstrate various properties of that creature (e.g., having six limbs—two human arms and four equine legs). How, then, does a thing that has a “true and immutable nature” differ from one that does not? Unless this question can be satisfactorily answered, Descartes’s theory of true and immutable natures cannot save his Ontological Argument from Gaunilo’s objection.²

Gaunilo’s objection, however, has a serious limitation: at best it shows that *something* is wrong with the argument. But it does not show *what* is wrong with it—that is, which premise is false or which step is fallacious. By contrast, the next objection to be considered attempts to pinpoint the error in the argument.

² Ernest Sosa has pointed out that Descartes both affirms and denies that a composite figure, such as a triangle inscribed in a square, has a “true and immutable nature” (quoted by Anthony Kenny, *Descartes*, p. 154; cf. CSM II 84, M 101, SPW 137–38, AT VII 117–18).

2.2 Kant's Objection

The most famous objection to the Ontological argument was made by Immanuel Kant in his *Critique of Pure Reason* (1781). Kant attacks a premise that we have seen to be common to both Anselm's and Descartes's arguments, namely, the premise that existence contributes to a thing's greatness, or that existence is a perfection. Kant attacks this premise in a particularly deep way, by identifying and criticizing an assumption behind it. This assumption is that existence is a *property*. The premise that existence is a perfection rests on this assumption, because a perfection is a particular type of property, namely, one that makes a thing better than it would be without that property. To see this, consider, for example, the perfections of God: he is said to be *omnipotent* (all-powerful), *omniscient* (all-knowing), and *omnibenevolent* (all-good). Clearly, each of these three perfections would be a characteristic, property or attribute of God. This illustrates the general principle that perfections are properties. Now the Ontological Argument holds that another of God's perfections is *existence*. By the same principle, then, it too must be a property. So we may take as established the statement

- (1) If existence is a perfection, then existence is a property.

Now Kant's strategy is to attack (1)'s consequent. If his attack succeeds in showing that (1)'s consequent is false (i.e., that existence is not a property) then, by *modus tollens*, it also shows that (1)'s antecedent is false, that is, that existence is not a perfection. This would refute Descartes's Ontological Argument, because, as we have seen, the subargument for the crucial part of the longer version's second premise, as well as the concise version itself, both use the statement that existence is a perfection as a premise.

But how does Kant propose to show that existence is not a property? Here is part of what he actually says:

"*Being*" is obviously not a real predicate; that is, it is not a concept of something which could be added to the concept of a thing. It is merely the positing of a thing, or of certain determinations, as existing in themselves . . . The proposition, "God is omnipotent," contains two concepts, each of which has its object—God and omnipotence . . . If, now, we take the subject (God) with all its predicates (among which is omnipotence), and say "God is," or

"There is a God," we attach no new predicate to the concept of God, but only posit the subject itself with all its predicates, and indeed posit it as being an object that stands in relation to my *concept*. The content of both must be one and the same; nothing can have been added to the concept, which expresses merely what is possible, by my thinking its object (through the expression "it is") as given absolutely. Otherwise stated, the real contains no more than the merely possible. A hundred real thalers do not contain the least coin more than a hundred possible thalers . . . By whatever and by however many predicates we may think a thing—even if we completely determine it—we do not make the least addition to the thing when we further declare that this thing *is*.³

Kant's thesis in this famous passage is his opening statement that "*Being*' is obviously not a real predicate." Putting this into contemporary terminology, Kant may be interpreted as saying

(2) "Exists" is not a descriptive predicate.

This claim, to which we shall return in a moment, has a direct bearing on whether existence is a property; for the statement would be generally accepted,

(3) If existence is a property, then "exists" is a descriptive predicate.

But, from (2) and (3), it follows by *modus tollens* that existence is not a property—in which case, as we have seen, it is not a perfection. Thus, if (2) and (3) are both correct, then the Ontological Argument is refuted.

Statement (3) is certainly acceptable; for a property (e.g., redness) is a characteristic, quality or feature that a thing may have or lack. So, a word that designates a property (e.g., "red") functions to describe things as having or lacking that property; it is a descriptive word, or, in more technical vocabulary, a descriptive predicate; for example, if red(ness) is a property, then the word "red" is a descriptive predicate. The general principle involved here—that if X is a property, then a word that designates

³ Immanuel Kant, *Critique of Pure Reason*, pp. 504–5.

X is a descriptive predicate—is an uncontroversial, analytic truth. Statement (3) simply applies this principle to existence.

Since (3) is unproblematic, the whole weight of Kant's objection rests on (2)—the contention that “exists” is not a descriptive predicate. Accordingly, we must consider exactly what Kant means by (2). He does not mean that the term “exist” can't take the predicate place in a sentence. This would be obviously false, because in a sentence like “Tame tigers exist,” the term “exist” is the grammatical predicate. (This is a case where the verb and predicate are telescoped into one word; the untelescoped subject-verb-predicate construction would read, “Tame tigers are existent things,” or the like.) But what Kant says is that “exist” is not a *descriptive* predicate (not a “*real* predicate”). Why? Well, compare “Tame tigers exist” with “Tame tigers growl.” Here “growl” is again the grammatical predicate. But it also functions to *describe* tame tigers, to say something about what they are like: they growl. On the other hand, the term “exist” in “tame tigers exist,” according to Kant, does not function to describe tame tigers. If I tell you that tame tigers exist, I have not told you anything about what tame tigers are like. This is why Kant says that “exists” is not a descriptive predicate. A descriptive predicate is one that tells what a thing is like. Since “exists” doesn't do this, it is merely a grammatical predicate, not a descriptive one. Thus Kant says that “being” is not a *real* predicate.

What is the true function of the term “exist” in a sentence like “Tame tigers exist”? According to Kant, its true function is to say that *the concept of a tame tiger applies to something*. So when I say that tame tigers exist, I am not saying that tame tigers have a certain property, namely, existence. I am saying something about the concept of a tame tiger. I am saying that this concept applies to something, or has instances, or is exemplified. Today, some philosophers who agree with Kant's basic point would put it in terms of language, rather than concepts. They would say that the function of “exist” in “Tame tigers exist” is to say that the *term* “tame tiger” applies to something.

The key to grasping Kant's point, whether it is put in terms of concepts or of language, is to see the contrast between (a) describing a thing and (b) saying that a concept or a term applies to something. A further example may help to make this contrast more vivid. Suppose I say “God is omnipotent.” Then I am describing God; I am telling you something about what he is like; I am ascribing a certain property or characteristic to him. But suppose I say “God exists.” Then I am not describing God; I am not

telling you anything about what he is like; I am not ascribing any property or characteristic to him. Instead, I am saying that the term “God,” unlike the term “unicorn,” applies to something; that the concept of God, unlike the concept of a unicorn, is exemplified or instantiated. Thus, there is a deep difference between statements of the form “X is of such-and-such a kind” and “X exists.” The former describe X, assign a property or properties to X; the latter do not. Instead, they covertly mention the concept of an X or the term “X”; they are equivalent to “the concept X has instances” or “the term ‘X’ applies to something.”

Kant’s view about existence—and with it his objection to the Ontological Argument—is quite widely accepted by contemporary philosophers. There are probably many philosophers who regard the proposition that existence is not a property as being about as well-established as any philosophical thesis can be. Kant’s view is even reflected in the symbolic notation of modern logic. A descriptive statement, like “the Taj Mahal is white,” would be symbolized as “Fa” (read as “a is F” or “a has the property F”), where “a” is a constant denoting the Taj Mahal and “F” is a predicate designating the property, whiteness. But an affirmative existential statement (i.e., a statement asserting that something exists), like “the Taj Mahal exists,” would be symbolized as “ $\exists x(x = a)$ ” (read as “there exists something x such that x is identical with the Taj Mahal”), where “ \exists ” is called the *existential quantifier* and “x” is a variable. The point, which can be appreciated without mastering the technicalities involved, is simply that existence is not represented as a predicate designating a property but by means of the quantifier together with the variable, “ $\exists x$ ” (read as “There exists something x such that . . .”). Indeed, “ $\exists x(x = a)$ ” contains no predicate expression at all.

Although many philosophers accept Kant’s criticism of the Ontological Argument, there are also philosophers who dispute it. These philosophers would point out that properties differ widely from each other (compare, e.g., the property of whiteness and the property of omnipotence). So why couldn’t existence be a property, even if a rather special one? Existence could be a property that such things as the Taj Mahal, Australia and electrons have and such things as Santa Claus, Shangri-La, and gremlins do not have. And “exists” could be a descriptive predicate used to designate this property. The fact that existence is not usually treated as a predicate in the symbolic notation of logic proves nothing; for this notation is only meant to facilitate the evaluation of arguments as valid or invalid and does not reveal any metaphysical truths. Besides,

even in this notation it is not impossible, but only inconvenient, to treat existence as a property. “The Taj Mahal exists,” for example, can be symbolized as $\exists x[(Ex) \& (x = a)]$ (read as “There is something x such that x exists and x is identical with the Taj Mahal”), where existence is represented by the predicate E instead of the quantifier and variable, $\exists x$. Therefore, some philosophers would say, Kant’s position is only one possible view of the matter. To show that it is the true view, argument is required.

As one would expect, a number of arguments have been offered in support of Kant’s view. But no such argument is generally regarded as conclusive. Kant’s own reasoning in the passage we quoted above, for example, is unlikely to sway a philosopher who sees nothing wrong with regarding existence as a property. Its core is contained in the following segment:

If . . . we say “God is,” or “There is a God,” we attach no new predicate to the concept of God, but only posit the subject itself with all its predicates, and indeed posit it as being an object that stands in relation to my *concept*. The content of both must be one and the same; nothing can have been added to the concept, which expresses merely what is possible, by my thinking its object (through the expression “it is”) as given absolutely. Otherwise stated, the real contains no more than the merely possible. A hundred real thalers [large silver coins formerly used throughout Europe] do not contain the least coin more than a hundred possible thalers . . . By whatever and by however many predicates we may think a thing—even if we completely determine it—we do not make the least addition to the thing when we further declare that this thing *is*.

Athough this passage may seem rhetorically quite powerful, its argument is difficult to make out. Its gist seems to be that if existence were a property, there would be a mismatch between our concept of God and the object of that concept; for then when we say or think that God exists, our concept of God would fail to include a property that we had attributed to God, namely existence. The reason why our concept of God would fail to include that property is that our concept of an object can never include existence, even if we say or think that the object really exists. Why? Because the concept of a real X cannot include any property that the

concept of a merely possible X does not include, and the concept of a merely possible X obviously does not include the property of existence. To put it more formally, perhaps we can fairly reconstruct Kant's argument this way:

- (1) If existence is a property, then when we say that God exists our concept of God matches its object only if our concept of God includes (the property of) existence.
- (2) If when we say that God exists our concept of God matches its object only if our concept of God includes (the property of) existence, then (likewise) when we say that one hundred real thalers (or any other real object; the hundred real thalers is only an example) exists our concept of one hundred real thalers matches its object only if our concept of one hundred real thalers includes (the property of) existence.
- (3) Our concept of one hundred real thalers does not include any property that our concept of one hundred (merely) possible thalers does not include.
- (4) If our concept of one hundred real thalers does not include any property that our concept of one hundred (merely) possible thalers does not include, then it is not the case that our concept of one hundred real thalers matches its object only if our concept of one hundred real thalers includes (the property of) existence.

∴ (5) Existence is not a property.

This is a logically valid argument.⁴ It may even be sound, and we venture to say that like many other Kantian arguments, it goes deep. The trouble with it, however, is that anyone who holds that existence is a property

⁴ For readers who know basic logic, here is a proof of its validity.

- (1) $p \supset (q \supset r)$
 - (2) $(q \supset r) \supset (s \supset t)$
 - (3) u
 - (4) $u \supset \sim(s \supset t)$
-
- _____
∴ $\sim p$
- (5) $\sim(s \supset t)$ (3), (4), *Modus Ponens*
 - (6) $\sim(q \supset r)$ (2), (5), *Modus Tollens*
 - (7) $\sim p$ (1), (6), *Modus Tollens*

would say that premise (3) just begs the question; for she would say that our concept of one hundred real thalers does include a property that our concept of one hundred (merely) possible thalers does not include, namely, the property of existence! Kant appears to think that he can refute that claim by pointing out that “a hundred real thalers do not contain the least coin more than a hundred possible thalers,” but this is irrelevant; for a person who holds that a hundred real thalers has the property of existence and that a hundred merely possible thalers lacks that property certainly does not mean that a hundred real thalers has the property of containing more coins than a hundred merely possible thalers. Rather, the person means that a hundred real thalers has a property that a hundred merely possible thalers lacks, namely, the property of existence. Likewise, to generalize the point, someone who holds that a real object with properties P₁, P₂, P₃ . . . P_n (where none of these is existence) also has a property that a merely possible object with properties P₁, P₂, P₃ . . . P_n lacks does not mean that the real object has some property P_x, *other than existence*, that the possible object lacks, but rather that the real object has a specific property that the merely possible one lacks: existence.

In the following section, we shall offer a defense of Kant’s objection. We shall not claim, however, that this defense is conclusive, but only that it makes a reasonable and illustrative case for Kant’s position.

2.3 Further Consideration of Kant’s Objection

The following defense of Kant’s objection is somewhat indirect. We shall begin by presenting a problem that at first seems quite unrelated to Kant’s objection. Then we shall argue that the most plausible solution to this problem commits one to Kant’s view about existence.⁵

Consider the statement

- (1) Carnivorous cows do not exist.

This kind of statement, which denies the existence of something, is called a “negative existential statement,” or just a “negative existential.” Negative existentials raise a simple yet perplexing problem that has puzzled

⁵ In presenting the problem and possible solutions, we shall draw on Richard Cartwright, “Negative Existentials,” 629–39.

philosophers at least since the time of Plato. This problem is, How can such a statement be true? To be true, it must be meaningful. But for the statement to be meaningful, it seems that its subject-term must pick out something, namely, carnivorous cows. But if the subject-term does pick out carnivorous cows, then carnivorous cows exist after all, so the statement that they don't exist must be false. For the same reason, it looks as though all negative existentials must be false—which is surely absurd.

There are two classic solutions to this problem: *Inflationism* and *Deflationism*. The purpose of both solutions is to show how negative existentials can be both meaningful and true.

Inflationism, which is defended in the early writings of Bertrand Russell and in the works of the Austrian philosopher Alexius Meinong (1853–1920), tries to allow for true negative existentials by drawing a distinction between *existence* and *being* or *subsistence*. The basic idea is that anything that can be thought about or talked about must have being, must subsist, even if it does not exist. As applied to our example, the idea is that although carnivorous cows do not exist, they do subsist or have being. So statement 1 really says,

(1a) Carnivorous cows subsist (have being), but don't exist.

In (1a), the subject-term does pick out something, namely, subsisting carnivorous cows. So the statement is meaningful. Moreover, it also happens to be true, since these subsisting carnivorous cows *only* subsist and do not exist. The upshot is that by interpreting (1) to mean (1a), (1) can be seen to be both meaningful and true.

Inflationism, however, is not a popular theory among contemporary philosophers, for it has at least two very troublesome consequences. First, the realm of “being” that Inflationism postulates is very mysterious. Suppose one asks, *Where* are the carnivorous cows and other entities that only subsist but don't exist? Well, the answer has to be that they are nowhere; for if they were somewhere (i.e., if they were situated in space) then they would exist. But if they are nowhere, what can it mean to say that they “subsist” or “have being”? The Inflationist view seems utterly mysterious.

The second difficulty is that when Inflationism is applied to some negative existentials, it is downright paradoxical. Consider, for example, “Round squares do not exist.” According to Inflationism, this statement really says that “round squares subsist (have being), but don't exist.” But

how can a round square, which is a contradiction in terms, subsist or have being? It would seem that such an entity—one that both has and does not have angles—cannot subsist or have being anymore than it can exist. It cannot have any kind of being at all.

Deflationism, which is defended in the mature writings of Bertrand Russell and favored by most contemporary philosophers, attempts to allow for true negative existentials without running into the paradoxes of Inflationism. The fundamental claim of Deflationism is that negative existentials *are not really about their subjects at all*. A sentence like “Carnivorous cows do not exist” may *look* as if it were about carnivorous cows, but this is just a misleading appearance created by its grammatical form. Since there are no carnivorous cows for the sentence to be about, it cannot really be about carnivorous cows.

Deflationists support this basic claim by offering paraphrases of negative existential statements that are meant to reveal their true meaning. Since these paraphrases are supposed to reveal or to analyze the statements’ true meaning, the paraphrases are called “analyses” of the statements. Deflationist analyses of negative existential statements differ somewhat from one Deflationist philosopher to another. Here are some typical Deflationist analyses of “Carnivorous cows do not exist”:

- (1b) The concept *carnivorous cow* does not apply to anything; i.e., this concept is not exemplified: it has no instances (Kant, Frege).
- (1c) The term “carnivorous cow” does not apply to anything.
- (1d) Nothing has the defining characteristics of a carnivorous cow; nothing combines the properties of being mammalian, bovine, and meat-eating. (C. D. Broad)⁶
- (1e) Nothing has the properties of being both carnivorous and a cow
there are no x 's such that “ x is carnivorous and x is a cow” is true.
(Bertrand Russell)⁷

⁶ “Suppose, e.g., that a ‘dragon’ is defined as a reptile which flies and breathes fire. Then the statement that dragons do not exist is equivalent to the statement that nothing combines the three properties of being a reptile, of flying, and of breathing fire” (C. D. Broad, “Arguments for the Existence of God,” www.ditext.com/broad/aeg.html).)

⁷ This is an extrapolation from what Russell says about “the whole realm of non-entities” in his famous essay “On Denoting,” p. 54, as reprinted in his *Logic and Knowledge: Essays 1901–1950*, pp. 41–56.

Contemporary philosophers regard these analyzes as being at least roughly equivalent to each other. One reason for this is that they would all be symbolized in the same way in modern logic: “ $\sim(\exists x)(Vx \& Cx)$ ” (read as: “It is not the case that there exists an x such that x is carnivorous and x is a cow”). For our purposes, the important point is that these Deflationist analyses do not postulate any objects that subsist but don’t exist. In light of the difficulties such objects generate, therefore, Deflationism appears to be a better solution to the problem of negative existentials than Inflationism.

How does this issue bear on Kant’s objection to the Ontological Argument? The answer is that Deflationism commits one to Kant’s view about existence. The reason for this is twofold: (1) if we analyze negative existentials in the Deflationist manner, then we must analyze affirmative existentials in a parallel or isomorphic manner; but (2) to analyze affirmative existentials in a manner parallel to the Deflationist analysis of negative existentials is to adopt Kant’s view that “exists” is not a descriptive predicate. Let us explain each of these two points in turn.

To understand point 1, consider a pair of affirmative and negative existential statements pertaining to the same thing, for example, “Unicorns exist” and “Unicorns do not exist.” Now these statements contradict each other; they are what logic calls “contradictories” (of each other). This means that they must have opposite truth-values: they cannot *both* be true, nor can they *both* be false. If one is true, the other is false; and vice-versa. Now the reason for (1) is that if we do not analyze the two statements in a parallel fashion, then they won’t contradict each other. To see this, suppose that we analyze the negative existential “Unicorns don’t exist” as “The concept *unicorn* is not exemplified” and that we analyze the corresponding affirmative existential, “Unicorns exist,” as a predicative statement, that is, as “Unicorns have the property of existence.” Then it can be shown that the two statements do not contradict each other.

To show this, we need to use three statements:

1. Unicorns have the property of existence.
2. The concept *unicorn* is not exemplified.
3. Unicorns subsist or have being.

The following argument proves that (1) and (2)—which must be contradictories if they are correct analyses of “Unicorns exist” and “Unicorns don’t exist,” respectively—are not contradictories. Since contradictories must have opposite truth-values, we know

(A) If (1) and (2) are contradictories, then not-(1) entails (2).

Now, our discussion of Inflationism has shown that some philosophers maintain that not-(1) entails (3). Without going that far, we can at least agree with inflationists that:

(B) Not-(1) does not entail not-(3).

Further, we may affirm

(C) (2) entails not-(3).

This is because if there were such things as subsisting unicorns, or unicorns that lack existence but have being, then *they* would exemplify the concept *unicorn*. Now, the statement is necessarily true,

(D) If not-(1) does not entail not-(3) and (2) does entail not-(3), then not-(1) does not entail (2).

This is because logical entailment is a transitive relationship: if *P* entails *Q* and *Q* entails *R*, then *P* entails *R*. So, if not-(1) did entail (2), then it would have to entail everything that (2) entails, including not-(3). Therefore, if not-(1) does not entail not-(3), then it cannot entail (2), either. But now, there follows from (B), (C), and (D)

(E) Not-(1) does not entail (2).

Finally, there follows from (A) and (E)

(F) (1) and (2) are not contradictories.

This proves that if we analyze “unicorns exist” as (1) and “unicorns do not exist” as (2), then these statements will not contradict each other—which is absurd.

How then should we analyze “Unicorns exist,” so as to make it contradict “The concept *unicorn* is not exemplified”? The answer is obvious. We must analyze it as: “The concept *unicorn* is exemplified.” But (and this is the second point that we set out to explain) to analyze “Unicorns exist” in this way is to adopt Kant’s view about existence; for it is to reject the idea that the sentence describes what unicorns are like, in favor of the

idea that the sentence says that the concept of a unicorn is exemplified, or that the term “unicorn” applies to something. Thus, it is to reject the view that “exists” is a descriptive predicate. And this, as we have seen, in turn implies that existence is not a property and hence not a perfection, so that the Ontological Argument is unsound. The upshot is that if Deflationism is correct, then the Ontological Argument is unsound.

Although the foregoing defense of Kant’s objection seems plausible, it is not wholly unproblematic. For Deflationism is not without its difficulties. The main difficulty arises from negative existentials that deal with mythological and fictional creatures. For example, consider the sentence, “Dragons do not exist.” According to Deflationism, this sentence is not really *about* dragons at all, since there are no dragons for it to be about. Rather, the sentence means “The concept of a dragon is not exemplified” or “the term ‘dragon’ does not apply to anything.” Now while such an analysis seems quite plausible when applied to “carnivorous cows do not exist,” it sounds somewhat paradoxical when applied to a sentence dealing with dragons. For dragons, unlike carnivorous cows, are mythological creatures; they have a place in mythic lore and literature and a certain “status,” so to speak. The same difficulty arises when Deflationism is applied to sentences about fictional creatures. For example, to say that “Hamlet did not exist” is not really about Hamlet sounds paradoxical. Now it may well be that this difficulty is not fatal to Deflationism but only calls for certain refinements in the theory, designed to deal with fictional discourse. Indeed, philosophers of language and of art pursue research on this very topic. But pending a satisfactory analysis of mythological and fictional discourse within the general framework of Deflationism, it must be admitted that the foregoing defense of Kant’s objection cannot be regarded as conclusive.

2.4 Caterus’s Objection

Kant’s objection to the Ontological Argument, as we have seen, raises complex and far-ranging issues. Yet, some philosophers feel that the Ontological Argument commits a rather simple mistake—one that should not require such elaborate argumentation, and even theorizing, to expose; for the argument tries to prove God’s existence from a mere definition of the word “God.” But, according to these philosophers, it is impossible to deduce the existence of anything from a mere definition; it is impossible to “build bridges from the conceptual to the real.” This

objection to the Ontological Argument is actually an old one. Aquinas raised it in the *Summa Theologica* (Part I, Question 2, art. 1); and it was very concisely and clearly stated by the author of the first set of *Objections* to Descartes's *Meditations*, a priest named Caterus. Caterus wrote:

Even if it is granted that a supremely perfect being carries the implication of existence in virtue of its very title, it still does not follow that the existence in question is anything actual in the real world; all that follows is that the concept of existence is inseparably linked to the concept of a supreme being. (CSM II 72, M 98–9, SPW 136, AT VII 99)

In this section, we shall argue that Caterus's objection (as we shall call this criticism) is a decisive objection to Descartes's version of the Ontological Argument.

In order to explain Caterus's objection, we shall use a contemporary distinction that was not known to Caterus but is implicit in what he wrote, namely, the distinction between the *material mode of speech* and the *formal mode of speech*. This distinction, as well as its bearing on the Ontological Argument, is introduced in a brilliant passage by the English philosopher, Antony Flew (1923–2010). The Ontological Argument, Flew writes, provides

a memorably persuasive illustration of the need to have and to master a notation which can make absolutely and systematically clear the fundamental difference between, on the one hand, verbal and conceptual questions and, on the other, “matters of fact and existence.” For the premise of the Ontological Argument is the definition of a word, whereas the proposed conclusion would be the supreme fact of the universe. The prime need is to distinguish, and the corresponding notational need is for devices to mark the distinction, between: on the one hand, discussion of the concept or concepts of—say—God (talk about the meaning or meanings of the word *God*, and about the implications of its employment); and, on the other hand, discussion of the objects, if any, of these concepts (talk about the things to which these words do or would refer) . . . Once given the prime distinction between concept and object, and a corresponding notation, these can be and have been developed

to illuminate . . . logical relations. The extended notational distinction is that between the Material mode of Speech (MMS) and the Formal Mode of Speech (FMS). These labels were introduced by Rudolf Carnap (b. 1891), a charter member of the Vienna Circle of old original logical positivists. . . . The man who says that Threehood necessarily involves Oddness, that the idea of a triangle contains the idea of the equality of its three angles to two right angles, or that existence is part of—or is—the essence of God is thereby employing the Material Mode of Speech. The alternative Formal Mode is, as might be feared much more long-winded: to say “There are three of them” and to deny “There are an odd number of them” would be to contradict yourself; to say “This is a triangle” and to deny “This has three angles equal to two right angles” would be to contradict yourself; and “which exists” is part of—or is—the definition of the word *God* . . . No one would be so foolish as to suggest that the Formal Mode ought completely and generally to replace the Material Mode of Speech. Yet it is extremely important to be able, and on occasion willing, to transpose passages from the one into the other. [An] excellent exercise would be . . . to transpose into the Formal Mode the passage from . . . Descartes quoted already [in which Descartes advances the Ontological Argument].⁸

In a moment, we will do the “exercise” that Flew proposes. First, however, let us make sure that we understand the distinction between the material and formal modes of speech. The material mode is the more common mode of discourse, in which people talk about objects in the world, like rocks and trees. The formal mode, on the other hand, is a more specialized mode, in which we talk *about concepts and/or words*. Here you may object that concepts or words are “in the world,” just as rocks and trees are. This is true. But there is still a distinction between language and concepts on the one hand, and nonlinguistic or nonconceptual reality on the other. Formal mode discourse has the former as its subject matter; material mode discourse has the latter as its subject-matter. For example, suppose that I say or write:

⁸Antony Flew, *An Introduction to Western Philosophy*, pp. 186–88.

(a) Horses are large, solid-hoofed, herbivorous quadrupeds.

This statement is in the material mode, because it is about certain “objects in the world” (i.e., certain nonlinguistic and nonconceptual items): horses. But suppose that instead of (a), I say,

(b) “Horse” means “large, solid-hoofed, herbivorous quadruped.”

Then my statement is in the formal mode, because it is about certain words or linguistic expressions: “horse” and “large, solid-hoofed, herbivorous quadruped.” Another formal-mode statement would be:

(c) The concept *horse* may be applied when and only when the concept *large, solid-hoofed, herbivorous quadruped* may be applied.

The reason (c) is in the formal mode, of course, is that it is about certain concepts: the concepts *horse* and *large, solid-hoofed, herbivorous quadruped*. It would still be in the formal mode if it were expanded to read

(c') The concept *horse*, or term “horse,” may be applied when and only when the concept *large, solid-hoofed, herbivorous quadruped*, or phrase “large, solid-hoofed, herbivorous quadruped,” may be applied.

Now as Flew points out, it would be very foolish to suggest that the formal mode is always preferable to the material mode. Indeed, if we had to transpose all our material mode statements into formal mode ones, then we simply could not say most of the things that we want to say; for we generally mean to talk about nonlinguistic reality, rather than about concepts or language itself. The two types of discourse are not equivalent, except perhaps in certain special instances, such as definitions, which are arguably about words rather than things. (This is the reason for talking of “transposing,” rather than “translating,” from one mode to the other. The notion of transposition is borrowed from music. Just as transposing a piece of music from one key into another may change its character, so transposing a statement from one mode into the other may alter its meaning.)

Nevertheless, there are certain contexts or situations where it is important to be able to transpose a material mode statement into a formal mode one. For example, suppose I say,

(d) A unicorn is a one-horned animal resembling a horse.

A person who did not know that a unicorn is a mythical beast might be misled by my statement into thinking that there really are unicorns or that I believe that there are unicorns; or a person who knew that there are no unicorns might begin to wonder how my statement can be meaningful and even true, since there are no unicorns for it to be about. In such cases, it would be helpful to transpose (d) out of the material mode into the formal mode, as follows:

(e) “Unicorn” means “one-horned animal resembling a horse.”

The transposition makes it clear that (d) is only a definition, and as such, does not imply that there are any objects actually answering to the definition. Like other definitions, (d) does not carry what philosophers call “existential import”: it does not imply the existence of anything. All that (d) implies is that we have a concept of a unicorn and a corresponding term. Transposing (d) into (e) makes this point obvious.

We are now ready to apply what we have learned about the material and formal modes to Descartes’s Ontological Argument. Let us begin with the very simple version of the argument that, as we saw in section 1, he offers in connection with the last of his three possible objections to the argument:

- (1) A supremely perfect being has all perfections.
- (2) Existence is a perfection.

. . . (3) A supremely perfect being exists.

To refute this argument, we shall make a series of four remarks.

1. Premise (1) is a definition; for although it is not explicitly couched in the form of a definition, it turns on the point that a *supremely perfect* being must, by definition, possess *every* perfection. Now since (1) is a definition, it does not carry existential import; it does not imply or presuppose the existence of anything. Specifically, (1), or the use of the noun phrase “a supremely perfect being” within (1), does not imply or presuppose the existence of a supremely perfect being. This is fortunate because if (1) *did* imply or presuppose the existence of a supremely perfect being, then the argument would of course beg the question; for then its first

premise would covertly assert the very proposition that the argument is intended to establish, so that it would be impossible to know that premise to be true without already knowing the conclusion to be true.

2. Since (1) is a definition, it can easily be transposed out of the material mode into the formal mode:

(1F) “Supremely perfect being” means “being that has all perfections.”

As already indicated, some philosophers would even say that transposing (1) into (1F) does not alter (1)’s meaning, because (1) is a definition and definitions are, strictly speaking, about words rather than things. But the criticism we are developing does not depend on whether this point is correct, so we need not insist on it.

3. A proponent of the argument must agree to let us substitute (1F) for (1) in the argument, on pain of begging the question; for why should he refuse to allow this substitution? The only possible reason would be that he wishes to treat (1) as *more* than just a definition—that he is interpreting (1), or the noun phrase “a supremely perfect being” in (1), as carrying existential import, that is, as implying the existence of a supremely perfect being. But then the argument begs the question, as we have seen.

4. However, when (1F) is substituted for (1), the argument becomes invalid. For (3) does not logically follow from (1F) and (2). The only conclusion that may be derivable from (1F) and (2) is

(3F) “Supremely perfect being” means “being that (among other things) exists.”

But this conclusion does not say that a supremely perfect being exists. It merely says that *only a being whose definition says that (among other things) it exists would satisfy the definition of a “supremely perfect being”*; or, as it might very misleadingly be put, that only a being that “exists by definition” would satisfy the definition of a “supremely perfect being.” But this does not show that any being does satisfy this definition; it does not prove that *there is* a supremely perfect being. Look at it this way: To be supremely perfect, a being would have to be omnipotent, omniscient, and omnibenevolent. Now how could the fact that only a being whose definition says that it exists would satisfy the definition of a supremely perfect being possibly prove that there really *is* an omnipotent, omniscient, and

omnibenevolent being? That fact is simply *irrelevant* to such a conclusion. To put the point still differently: (3F) is, so to speak, purely “negative information”: it just means that no being whose very definition does *not* say (imply) that it exists qualifies for the title, “supremely perfect being.” But it doesn’t follow from this that any being does qualify for this title: perhaps the title simply has no “holder.”

At this point, you should reread the passage from Caterus quoted at the beginning of this section. Can you see how the criticism just offered merely spells out Caterus’s point more fully? All we have done is to use the contemporary distinction between the material and formal modes to explicate Caterus’s insight; for Caterus can be seen as pointing out that the formal mode statement “A supremely perfect being carries the implication of its existence by its very title” (= statement 3F) does not entail the material mode statement that “The existence in question is anything actual in the real world” (= statement 3) but only another formal-mode statement to the effect that the *concept* of existence is inseparable from the *concept* of a supreme being. Notice also that Caterus’s objection is completely independent of Kant’s. Caterus’s refutation works even if one grants the premise that existence is a perfection; for granting this premise merely compels one to define the concept of a supremely perfect being in such a way that only a being whose definition says that it exists can satisfy the definition. But, as Caterus shows, this does not mean that anything does satisfy the definition.

In his replies to the First Set of *Objections*, Descartes attempts to answer Caterus’s objection by making a distinction between *possible existence* and *necessary existence*. He writes:

It must be noted that possible existence is contained in the concept or idea of anything that we clearly and distinctly understand; but in no case is necessary existence so contained, except in the case of the idea of God. Those who carefully attend to this difference between the idea of God and every other idea will undoubtedly perceive that even though our understanding of other things always involves understanding them as if they were existing things, it does not follow that they do exist, but merely that they are capable of existing. For our understanding does not show us that it is necessary for actual existence to be conjoined with their other properties. But, from the fact that we understand that actual existence is necessarily and always contained

with the other attributes of God, it certainly does follow that God exists. (CSM II 83, M 100, SPW 137, AT VII 116–117)

In replying to Caterus, Descartes also offers an explanation of his notion of necessary existence. Unlike contemporary philosophers, who think of necessary existence as existence in all possible worlds, Descartes thinks of it as existence by virtue of a being's own power. Thus he writes:

When we attend to the immense power of this being, we shall be unable to think of its existence as [even] possible without also recognizing that it can exist by its own power; and we shall infer from this that this being does really exist and has existed from eternity, since it is quite evident by the natural light that what can exist by its own power always exists. So we shall come to understand that necessary existence is contained in the idea of a supremely powerful being, not by any fiction of the intellect, but because it belongs to the true and immutable nature of such a being that it exists. (CSM II 85, M 102, SPW 139, AT VII 119)

Descartes also alludes to necessary existence in two footnotes that he added in the French version of *Meditation V*. Since the French version of the *Meditations* appeared in 1647, six years after the Latin version appeared along with (all but the seventh set of) the *Objections and Replies*, it seems likely that Descartes added these two footnotes with his reply to Caterus in mind. In them, he expands the segment of the Latin version that Cottingham translates as “apart from God, there is nothing of which I am capable of thinking such that existence belongs to its essence” into “apart from God, there is nothing of which I am capable of thinking such that existence necessarily belongs to its essence,” and he expands the translated Latin version’s “what is more self-evident than the fact that the supreme being exists, or that God, to whose essence alone existence belongs, exists?” into “what is more self-evident than to think that there is a God, that is to say, a sovereign and perfect being, in the idea of whom alone necessary or eternal existence is comprised, and who consequently exists? (CSM II 47, SPW 47, AT VII 68–69).⁹

⁹ I have slightly amended Cottingham's translation of the second-quoted passage from the French version.

Does Descartes's appeal to the notion of necessary existence help to answer Caterus's objection? It seems not; for, first, the curious principle that "what can exist by its own power always exists," contrary to Descartes's claim that "it is quite evident by the natural light," is anything but obvious: compare it to "if I am thinking, then I exist" or to " $2 + 2 = 4$." Indeed, this principle seems to mean, or at least to imply, that just because we have the concept of something that can exist by its own power, therefore this concept must be instantiated, which is surely false. After all, just because we can form the concept of something that exists without any *external* cause, it does not follow that that concept is instantiated; but if we now add to the content of our concept of this thing that it exists because of an *internal* cause (or by its own power, of because it is the "cause of itself"), then even assuming—what might be questioned—that this addition makes sense, it still does not follow that the concept is instantiated. Second, Caterus could easily amend his objection, quoted above, to read this way:

Even if it is granted that a supremely perfect being carries the implication of *necessary* existence in virtue of its very title, it still does not follow that the existence in question is anything actual in the real world; all that follows is that the concept of *necessary* existence is inseparably linked to the concept of a supreme being. (CSM II 72, M 98–9, SPW 136, AT VII 99, with "necessary" added)

In short, just because one builds necessary existence into the concept of a supremely perfect being, it does not follow that anything answers to that concept. Thus, suppose that we import necessary existence into the formal-mode version of Descartes's shortest version of the Ontological Argument, so that it reads this way:

(1F) "Supremely perfect being" means "being that has all perfections."

(2N) Necessary existence is a perfection.

. . . (3FN) "Supremely perfect being" means "being that (among other things) necessarily exists."

This argument still does not show that a supremely perfect being necessarily exists, but only that a being whose definition says that (among other things) it necessarily exists would satisfy the definition of a

supremely perfect being. On the other hand, if we insist on keeping the argument in the material mode, then as before it begs the question.

Caterus's objection also applies to the version of the Ontological Argument that Descartes gives in answering his second hypothetical objection. That version went as follows:

- (i) I cannot think of God without existence.
- (ii) Existence is inseparable from God. (from step i)
- (iii) God really exists. (from step ii)

The objection is then that (ii) does not really follow from (i). All that really follows from (i) is:

- (iiF) The concept of existence is inseparable from the concept of God,

where this means that one can't think of God without thinking of existence (not that one can't think of existence without thinking of God). But even if (iiF) is true, this does not show that *there is* anything answering to the concept of God, or that this concept is exemplified. Again, the basic point is that just because we include existence in a concept, it does not follow that that concept is exemplified or instantiated—i.e., that there is anything answering to the concept. Nor would it help to substitute “necessary existence” for “existence” in (i), (ii), and (iiF) and “necessarily exists” for “exists” in (iii).

Before concluding our examination of the Ontological Argument, we should ask whether Caterus's refutation also works against Descartes's initial, longer formulation of the argument. That formulation was:

- (1) Whatever I clearly and distinctly perceive to belong to the nature or essence of a thing can be truly affirmed of that thing.
 - (2) I clearly and distinctly perceive that existence belongs to the nature or essence of a supremely perfect being.
-
- .: (3) Existence can be truly affirmed of a supremely perfect being, i.e., a supremely perfect being exists.

To see that Caterus's objection does indeed bear on this argument, we need only transpose it out of the material mode into the formal mode. One way to do this is as follows:

- (1F) Whatever I clearly and distinctly perceive to be part of the definition of “X” can be truly affirmed of all X’s, if there are any.
- (2F) I clearly and distinctly perceive that “exists” is part of the definition of “supremely perfect being.”

. . . (3F) “Exists” can be truly affirmed of all supremely perfect beings, if there are any.

This argument obviously does not prove that *there are* any supremely perfect beings or (more to the point) that there *is* even one supremely perfect being. Rather, all it proves is that *if* there are any supremely perfect beings, then they exist or (more to the point) that if there is a supremely perfect being, then it exists. In other words, the argument does not prove that God exists, but only (the tautology that) if God exists, then he exists. Notice that this point continues to hold if the phrase, “*if there are any*” is deleted from the argument: this phrase is used only for the sake of emphasis. Notice also that again, nothing would be gained by substituting “exists necessarily” (or “exists by virtue of its own power” or “exists in all possible worlds”) for “exists” in the argument. At best, the argument might then show that *if* God existed in the actual world, then he would exist in all possible worlds.¹⁰

It might be objected that this way of dealing with Descartes’s argument is too quick. For why should Descartes not simply refuse to allow the substitution of the formal mode argument for his material mode one? In response, we may make two points.

1. The formal mode version is an improved formulation of what was confusedly and tendentiously expressed in Descartes’s formulation. For, to quote Flew once again,

[A]n inferior notation may . . . encourage and express actually erroneous ideas. This is the reason for writing transpose rather than translate; for the FMS analogue may sometimes be a substantial improvement on, and hence not equivalent to, the MMS original. Thus . . . most of those, from Aristotle onwards, who have spoken of the essences of things would have been reluctant to allow that all they were saying was expressed in some FMS

¹⁰ This last observation is due to Krasimira Filcheva.

statement about the definitions of the words; though they might have had even greater difficulty in explaining precisely what more they had in mind, and in justifying their beliefs about it.¹¹

2. We need not alter Descartes's formulation as drastically as we have done in order to make Caterus's objection. Even the following formulation, in which only premise (1) is altered (and quite minimally so), will serve:

- (1Fa) Whatever I clearly and distinctly perceive to belong to the essence of an X can be truly affirmed of all X's, if there are any.
- (2) I clearly and distinctly perceive that existence belongs to the essence of a supremely perfect being.

.: (3Fa) Existence can be truly affirmed of all supremely perfect beings, if there are any.

Again, this argument obviously does not prove either that there are any supremely perfect beings or (more to the point) that there is even one supremely perfect being. Again, it yields only the tautologies that if there are supremely perfect beings, then they exist, or that that if there is a supremely perfect being, then it exists. Again, this point holds even if the phrase, "if there are any," is deleted from the argument. Finally, again, importing the notion of necessary existence into the argument would not help.

But what if Descartes refused to allow the substitution of (1F) or (1Fa) for (1)? Then our reply would be that substituting (1F) or (1Fa) for (1) makes it clear that not all things whose definitions (or "essences") we clearly and distinctly perceive to include certain properties must really exist. By contrast, refusing to substitute (1F) or (1Fa) for (1) amounts to insisting that all things whose definitions (or "essences") we clearly and distinctly perceive to include certain properties must really exist. But this would be obviously false. Furthermore, Descartes himself did not believe it; for in presenting his theory about "true and immutable natures" near the beginning of *Meditation V*, he explicitly says that many things whose true and immutable natures or essences he clearly and

¹¹ Antony Flew, *An Introduction to Western Philosophy*, p. 188.

distinctly perceives to include various properties “may not exist anywhere outside me.” For example, he says, a geometrical figure like a triangle has a true and immutable nature, whose constitutive properties he “clearly recognize[s],” “even if perhaps no such figure exists, or has ever existed, anywhere outside my thought” (CSM II 4, SPW 106, AT VII 64).

3. Some Implications for Descartes’s System

We may conclude that Descartes’s *Meditation V* argument for the existence of God is no more successful than his *Meditation III* arguments for God’s existence. This negative verdict on Descartes’s philosophical theology has major implications for the rest of his system. As we have said, his overall strategy, as shown in Figure 5.1, is to extract his criterion of truth from the *cogito*, to provide a vindication of this criterion by means of the *Meditation III* causal proofs of God’s existence, and then to use the criterion in his *Meditation V* ontological argument for God’s existence, in his *Meditation VI* attempt to establish that mind is a different substance than any matter which may exist, and finally in his attempt to prove, also in *Meditation VI*, that matter really does exist.

The first implication of the failure of Descartes’s arguments for the existence of God, then, is simply that his attempt to establish God’s existence must be judged unsuccessful. In view of the human significance of the question of God’s existence, and given that the subtitle of the *Meditations* includes the words, “in which are demonstrated the existence of God,” this, of course, represents a major failure in Descartes’s overall scheme.

The second implication is that Descartes’s attempt to vindicate his criterion of truth by appealing to the existence of a perfect God must be also be regarded as a failure. If the solution to the problem of the Cartesian

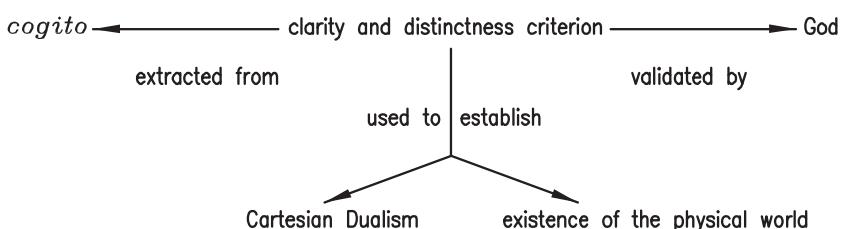


Figure 5.1

circle offered in the chapter 3 is satisfactory, then this failure does not stem from any circularity in Descartes's procedure but simply from the fact that his specific arguments for God's existence do not work. As we have noted, the *Meditation V* argument seems not to be intended to contribute to the vindication of Descartes's criterion of truth anyway. Nevertheless, if it were a sound argument, then perhaps it could play such a role provided the appeal it makes to clear and distinct perception were understood in the same way as for the *Meditation III* arguments. However, in light of the failure of both the *Meditation III* and the *Meditation V* arguments, it seems that we must conclude that Descartes's attempt to provide a divine vindication for his criterion of truth is ultimately unsuccessful. Therefore, if we hold that Descartes cannot legitimately use this criterion unless it can be vindicated, then we must also conclude that he cannot advance beyond the *cogito*. In particular, he cannot establish the two remaining major theses of his *Meditations*, the distinction between mind and (any) body (which may exist) and the existence of the material world, because his arguments for both of them rely on his criterion of truth.

Descartes's arguments for these two theses, however, are worth considering regardless of whether the criterion of truth that they rely upon can be vindicated. So in the next chapter, we shall adopt the following policy. We shall assume that the clarity-and-distinctness criterion of truth can stand on its own merits, or solely on the basis of the extraction argument (from the *cogito*) that was presented in chapter 3, section 1. In the course of examining Descartes's arguments for the distinction between mind and body and for the existence of matter, however, we shall inquire whether the failure of his philosophical theology has any implications for those arguments, beyond the fact that the criterion of truth that they employ must stand on its own merits. Our suggestion will be that the failure does not seriously compromise Descartes's argument for the distinction between mind and body, but that it profoundly affects his argument for the existence of matter.

Meditation VI

Dualism and the Material World

1. The Scope of *Meditation VI*

Meditation VI is a longish text in which Descartes completes his basic “agenda” and addresses a number of subsidiary themes. It can be divided into the following segments or episodes (the paragraph numbers are those of the Cottingham translation): (1) a discussion of the imagination and of how this faculty makes it “probable” that material things exist (paragraphs 1–3); (2) a review of what Descartes initially believed about the senses and of the reasons he subsequently found for doubting them (paragraphs 4–7); (3) the proof of the Real Distinction between mind and body (i.e., of Cartesian Dualism) (paragraph 9); (4) the proof that the material world exists (paragraph 10); (5) an account of certain particular beliefs about material things that, in light of this proof, can safely be accepted (paragraphs 11–14); (6) an account of certain plausible but nevertheless erroneous beliefs about material things (paragraph 15); (7) a physiological explanation of why we sometimes pursue ends that are bad for us and of why this fact does not reflect adversely on God’s perfection (paragraphs 16–23); and (8) a resolution of the dream problem (paragraph 24). In addition, episode 7 contains a subsidiary argument for mind-body dualism (paragraph 19). We shall not try to discuss all of these topics. Rather, we shall focus on Cartesian Dualism and on Descartes’s theory of the physical world. We shall examine Descartes’s main proof of dualism, his proof that material things exist, and some of his basic views about their nature; discuss the most famous problem that arises from Descartes’s dualism; and, finally, offer an overall assessment of Cartesian Dualism.

2. Descartes's Proof of the Real Distinction between Mind and Body

Descartes presents his main argument for dualism in the following paragraph, which we have divided into three segments in order to facilitate discussion:

[A] First, I know that everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it. Hence the fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct, since they are capable of being separated, at least by God. The question of what kind of power is required to bring about such a separation does not affect the judgement that the two things are distinct.

[B] Thus, simply by knowing that I exist and seeing at the same time that absolutely nothing else belongs to my nature or essence except that I am a thinking thing, I can infer correctly that my essence consists solely in the fact that I am a thinking thing.

[C] It is true that I may have (or, to anticipate, that I certainly have) a body that is very closely joined to me. But nevertheless, on the one hand I have a clear and distinct idea of myself, insofar as I am simply a thinking, non-extended thing; and on the other hand I have a distinct idea of body, insofar as this is simply an extended, non-thinking thing. And accordingly, it is certain that I am really distinct from my body, and can exist without it. (CSM II 54, SPW 114–115, AT VII 78)

Let us begin by examining segment B, since it raises special difficulties that need to be noticed before we can properly analyze Descartes's argument. On the face of it, segment B seems to be arguing

(i) I know that nothing belongs to my nature or essence except that I am a thinking thing.

.: (ii) My essence consists solely in the fact that I am a thinking thing.

Now in light of what philosophers call the “truth-condition for knowledge,” this is obviously a valid argument; for that condition says that, simply by virtue of the definition of knowledge, “S knows that *p*” entails that *p* is true. But still the argument is quite unsatisfactory; for nowhere in his previous *Meditations* has Descartes established the truth of its premise. Rather, all he has established (in *Meditation II*) is that the only property he *knows* for certain belongs to his nature or essence is thinking. This suggests that the argument in segment B really ought to go as follows:

(ia) I do not know that anything other than thinking belongs to my nature or essence.

. . . (ii) My essence consists solely in the fact that I am a thinking thing.

Indeed, in the French translation of the *Meditations*, which Descartes himself approved, this is essentially how the argument does go. The French version of segment [B] says,

Thus, just because I know with certainty that I exist, and that meanwhile I do not notice that anything else necessarily belongs to my nature or essence except that I am a thinking thing, I rightly conclude that my essence consists solely in the fact that I am a thinking thing.

As E. M. Curley has pointed out, the Latin version (on which the translation by John Cottingham from which we are quoting is based) is ambiguous and could also have been translated in this way.¹ We suggest that this would actually have been a better way for Cottingham to translate segment B. For it cannot be overemphasized that premise (ia), not premise (i), is all that Descartes is entitled to assert at this point in his *Meditations*: He showed, in *Meditation II*, that the only property which he knew for certain to belong to his essence was thinking; he did not show (but is now, in *Meditation VI*, trying to show) that the only property which in fact belongs to his essence is thinking. Therefore, the argument in segment B

¹ Edwin M. Curley, *Descartes Against the Skeptics*, p. 196.

should be interpreted, despite Cottingham's translation, as going from (ia) to (ii), not as going from (i) to (ii).

This observation, however, brings us face to face with the basic difficulty in the segment B argument: it is simply not valid. One cannot validly go from the premise that "P is the only property that is certainly known to belong to one's essence," to the conclusion that "P is the only property that really belongs to one's essence." Such an "argument from ignorance" is not significantly different from the obviously invalid "argument from doubt" that, as we saw, Descartes seems to use in his *Discourse on the Method* and *Search After Truth*: "I cannot doubt that I exist; I can doubt that (my) body exists; therefore I am not a body."

As we saw, however, in the more careful *Second Meditation*, Descartes admitted that this is an invalid argument; for directly after stating it, he went on to admit that, for all he knew, he might be a body (see page 89):

And yet may it not perhaps be the case that these very things which I am supposing to be nothing, because they are unknown to me, are in reality identical with the "I" of which I am aware? I do not know, and for the moment I shall not argue the point, since I can make judgements only about things which are known to me. (CSM II 18, SPW 82, AT VII 27)

Yet, as the word "thus" at the beginning of segment B indicates, Descartes now thinks that the material in segment A somehow *legitimizes* the reasoning in [B]. The material in [C], as we shall argue in a moment, also contributes to his legitimization of [B]. (More obviously, segment C spells out the consequence that follows when the general principles in [A] are combined with the thesis Descartes will soon prove [i.e., that he has a body] and with the clear and distinct conceptions of mind and body attained in his previous *Meditations*, namely, that his body is really distinct from his mind.) In any case, the paragraph as a whole is partly intended to show that an argument that was deemed unsatisfactory at an earlier stage, when supplemented by further principles, can be safely accepted. This reading is confirmed by what Descartes says in the *Meditations'* "Preface to the Reader." There he notes that it had been objected, against the *Discourse on the Method* that

From the fact that the human mind, when directed toward itself, does not perceive itself to be anything other than a thinking

thing, it does not follow that its nature or essence consists only in its being a thinking thing, where the word “only” excludes everything else that could be said to belong to the nature of the soul. (CSM II 7, AT VII 7–8)

After complaining that this objection is based on a misinterpretation of his intentions in the *Discourse*, Descartes adds, significantly,

I shall, however, show below how it follows from the fact that I am aware of nothing else belonging to my essence, that nothing else does in fact belong to it. (CSM II 7, AT VII 8)

This certainly raises a question of interpretation; for if an argument is invalid, then no magic can transform *that very argument* into a valid one. At best, the addition of new premises can only yield a new argument that includes the premises of the old one and is valid. But segments A and C do not seem to stand even in that relation to the argument in [B]. So what does Descartes mean by suggesting that the reasoning in [B] has now been vindicated? To answer this question, we must first examine what Descartes says in segments A and C.

The first sentence in [A] can be formulated as follows:

- (1) If I can clearly and distinctly conceive X existing in a certain way, then X can really exist in that way, at least by God’s power.

This opening premise, which Descartes also asserts at the very beginning of *Meditation VI* (where he says, “there is no doubt that God is capable of creating everything that I am capable of perceiving in this manner [i.e., clearly and distinctly]” [CSM II 50, SPW 110, AT VII 71], follows directly from his clarity-and-distinctness criterion of truth). In the next sentence, Descartes derives a general principle from (1)—that if he can clearly and distinctly conceive X existing apart from Y, then X is really a different thing from Y. (In expressing this principle we use the variables “X” and “Y” instead of Descartes’s expression “one thing apart from another,” because that expression may misleadingly suggest that in the if-clause of the principle, Descartes is already assuming that he is referring to two different things; whereas, in fact, the if-clause sets forth the condition under which we can know that we are referring to two different things rather than to only one, or that the variables “X” and “Y” stand for two different things

rather than for one and the same thing—i.e., the condition under which we can know that the then-clause is true.) The inference from (1) to this principle, however, requires two intermediate steps. The first one is elliptically stated in the clause, “since they are capable of being separated, at least by God.” Untelescopied, this is a premise saying,

- (2) If I can clearly and distinctly conceive X existing apart from Y, then X really can exist without Y, at least by God’s power.

This follows directly from (1) (because “existing apart from Y” counts as “existing in a certain way”), and leads to Descartes’s principle. To obtain that principle, however, the second intermediate step is required. This is a premise that Descartes states, somewhat obliquely, in his next sentence: “The question of what kind of power is required to bring about such a separation does not affect the judgment that the two things are distinct.” This can be put as follows:

- (3) If X really can exist without Y, no matter what power it takes, then X and Y are really two different things.

The idea behind (3) is that if X can really exist without or apart from Y, then *even if it takes so much as God’s power for this to happen*, X and Y must be different things; for not even God could make a thing exist without or apart from itself, since that is not even logically possible. (Here we are deliberately ignoring the extraordinary view about omnipotence that, as we saw in chapter 3, Descartes apparently held, according to which an omnipotent God could do logically impossible things—for example, make a four-sided triangle or make two contradictory statements both true. Descartes does not mention this mind-boggling view in the *Meditations*, and it would certainly ruin his argument for dualism. Indeed, he seems to put this view aside at the start of *Meditation VI* when he says, “I have never judged that something could not be made by him [God] except on the grounds that there would be a contradiction in my perceiving it distinctly” (CSM II 50, SPW 110–111, AT VII 71). From (2) and (3), we can now derive Descartes’s principle that:

- (4) If I can clearly and distinctly conceive X existing apart from Y, then X and Y are really two different things.

This brings us to the end of the material in segment [A].

The references to clearness and distinctness found in (1)–(4) are absent from the next segment of our passage, [B]. But they are picked up again in the following segment, [C], where Descartes advances another premise and then draws two conclusions. The premise, worded so as to make it connect clearly with (2) and (4), is

- (5) I can clearly and distinctly conceive myself, as a thinking and nonextended thing, existing apart from (my) body, as an extended and nonthinking thing.

This premise rests squarely on Descartes's long and careful discussion of the self in *Meditation II*; for surely, if anything was established by that discussion, it was that Descartes could form a clear and distinct conception of himself as simply a “thing that thinks” (affirms, denies, wills, imagines, seems to perceive, etc.)—a conception which, in the passage about the wax, he went on to contrast with the conception of an extended thing. But with the help of (5), the two conclusions that Descartes draws at the end of segment [C] can be derived. For there follows from (4) and (5)

- (6) I am really a different thing from my body.

And there follows from (2) and (5)

- (7) I can really exist without my body, at least by God's power.

These two conclusions assert the Real Distinction between mind and body. In his *Principles of Philosophy*, Descartes explains that a real distinction is one between two or more substances; he contrasts such a distinction with a modal distinction, which is one between a substance and a mode or between two modes of the same substance, and with a conceptual distinction or “distinction of reason,” which is one between a substance and an attribute without which it cannot be intelligibly conceived (he gives a substance and its duration as an example) or two such attributes of the same substance (he gives extension and being divided into parts as an example) (CSM I 213–214, SPW 180–181, AT VIIIA 28–29). Notice that despite the high profile of (4) in Descartes's text, this principle is not really needed in the argument; for it is not used to derive (7), and (6) can be derived without it from (2), (3), and (5). Indeed, as we shall

stress in section 6, lines (2), (3), and (5) are the argument's fundamental premises (see p. 315.)

Before discussing the argument, let us try to answer the question of interpretation raised above: How is this argument, which is extracted only from segments A and C, and which is essentially an argument from the independent conceivability of mind and body, supposed to legitimize the reasoning in segment B, which is an epistemological argument invoking the certainty of thought and (implicitly) the doubtfulness of body? Well, compare steps (5)–(7) with [B]. In (5)–(7), Descartes goes from the premise that he can clearly and distinctly conceive himself existing as only a thinking thing to the conclusion that he is distinct from his body and could exist without it. This certainly *resembles* the argument in [B], where he goes from the premise that thinking is the only property which he knows for certain belongs to his essence, to the conclusion that he is essentially only a thinking thing. But we can go further: we can say that the argument in [B] is merely a simplified version—which Descartes does not claim to be valid as it stands—of the argument extracted from [A] and [C]. To see this, let us compare the two arguments. In the first place, their *premises* are intimately related; for Descartes's claim in *Meditation II* that thought was the only property that he knew for certain belonged to his essence already had much more content than the words "Thinking is the only property that I know for certain belongs to my essence" reflect. What more? Well, he had a clear conception of his thinking; he was forming a clear conception of body; he could appreciate the contrast between those two conceptions; and he could clearly conceive, through the doubt, the possibility that the object of the former conception might exist though the object of the latter did not. But this is virtually what (5) says. In the second place, the two arguments' *conclusions* are intimately related; for, as Descartes uses these notions, to say that X can exist without Y entails that Y is not part of X's essence: "[I]f something can exist without some attribute, then it seems to me that that attribute is not included in its essence" (fourth set of *Replies*, CSM II 155, SPW 144, AT VII 219). Thus, when Descartes affirms in (7) that he can exist without his body, this entails that extension is not part of his essence, which leaves only thinking as his essence. But this is just the conclusion of [B].

Of course, the argument in [B] directly corresponds only to steps (5)–(7) of the argument in [A] and [C]. The longer argument goes beyond the short one, by presenting the general principles that are needed to legitimize the inference from a clearly and distinctly conceived distinction

between mind and body to a real distinction between them, namely, (1)–(4). But this should not stop us from seeing that the argument in [B] is a highly simplified version of what Descartes fully and adequately presents in (1)–(7).

This interpretation can be confirmed by three observations. First, the conclusion-indicator word “thus” that introduces segment B functions in a very unusual way. It does not function, in the normal way, to introduce the conclusion of an argument. Rather, it introduces *an entire argument*, namely, the segment B argument. Clearly then, Descartes does think that the longer argument somehow legitimizes the shorter one. But how? In the remark we quoted above from the “Preface to the Reader” Descartes seems to concede—and in *Meditation II* he plainly does concede—that the segment B argument is invalid. Presumably, then, Descartes is not saying that the argument has now been shown to be valid after all. Yet in the “Preface to the Reader” Descartes also promises to “show below” how such an argument can nonetheless be legitimate. The place where he fulfills this promise is the very paragraph in *Meditation VI* that we are examining. But surely his way of fulfilling the promise cannot be just to assert baldly that the argument earlier conceded to be invalid is really valid. What then is the meaning of Descartes’s “thus”? The answer, we suggest, is that this word indicates that the simplified, heuristic sketch in [B] is now legitimate, *because it serves as a stand-in for the more complex argument in [A] and [C]*. In other words, “thus” indicates that segments A and C provide a kind of second-order commentary on segment B, to the effect that the reasoning in [B] is now legitimate not because it is itself valid but rather because it can do duty for the more complex argument that Descartes has just presented.

Second, as Jakko Hintikka emphasized in his famous article on the *Cogito*, Cartesian doubt is no merely passive affair. Rather (as Descartes’s “litmus test” of certainty shows), “it amounts to an active attempt to think the contrary of what we usually believe.”² Thus, when Descartes says that he can be certain of his existence as a thinking thing while doubting the existence of his body, part of what this *means* is that he can clearly and distinctly conceive the possibility that his mind exists while his body does not. It should not be surprising, therefore, that Descartes initially presents an argument from the independent conceivability of

² Jakko Hintikka, “Cogito, Ergo Sum: Inference or Performance?” p. 124; also in Alexander Sesonske and Noel Fleming, eds., *Meta-Meditations: Studies in Descartes*, p. 63.

mind and body in the guise of an epistemological argument turning on certainty about the mind and doubt about the body.

Finally, when Descartes responded to a request that he present his main arguments “in geometrical fashion,” he himself gave, as his demonstration of the Real Distinction, essentially the argument in [A] and [C], omitting the one in [B] altogether (CSM II 119, SPW 158–9, AT VII 169–70). We may reasonably conclude, then, that the argument we have extracted from segments A and B is Descartes’s finished, “official” argument for dualism.

It behooves us, therefore, to understand exactly what this argument is, and is not, supposed to prove. It is supposed to prove, in conclusions 6 and 7, what Descartes calls “the Real Distinction between Mind and Body.” But exactly what does this distinction amount to? Well, Descartes is not claiming to have proved in (7) that his mind *does*, at any time, exist without his body (i.e., without the body that he will shortly argue he possesses). Rather, he is claiming to have proved that the mind *could* exist without the body. As line 3 indicates, Descartes holds that the mere possibility that mind can exist apart from body is enough to establish that they are different. As previously mentioned, his idea is that even if it would take so much as God’s power for this possibility to be actualized, the mere logical possibility is sufficient for mind and body to be different things; since a thing cannot possibly exist apart from itself. And (as he notes at the end of segment A) whether it would require God’s power for them to exist separately does not matter. If *any* power would allow them to exist separately, then they must be different. Thus, as the late, eminent Descartes scholar Margaret Wilson pointed out, it would be a fundamental misunderstanding to complain that

Descartes’s argument can show at best that mind and body are possibly or potentially distinct (would be distinct if God should choose to separate them)—not that they *are* distinct. [For] Descartes holds that two things *are* really distinct if it is *possible* for them to exist in separation. On his view actual distinctness does not entail *separateness*.³

Notice, then, that Descartes’s argument is not intended to establish that the mind or soul actually exists after the body stops functioning, that is,

³ Margaret Wilson, *Descartes*, p. 190.

that the immortality of the soul is a fact. Rather, it is only intended to show that the mind *could* exist without the body, that is, that immortality is *possible*. Notice also, however, that if the mind is just (some part of) the body, then immortality is not even possible—unless, of course, bodily resurrection is possible, a question we shall not discuss here. Thus, while Descartes's dualism is not *sufficient* to establish immortality, it does seem to be *necessary* (assuming that bodily resurrection is not possible).

Now that we have Descartes's main argument for dualism before us, it is time to evaluate it. Let us start by addressing a possible objection that has probably occurred to you already. This is that since the premises of Descartes's argument refer to God's power and since his arguments for God's existence are unsuccessful, his argument for dualism collapses. Although this objection is a very natural one to make, it does not really refute Descartes's argument; for the statements in the argument that refer to God's power, namely, (1), (2), and (7), need not depend for their *truth* on whether God exercises that power or even on whether God exists. To see this, consider premise 1, "If I can clearly and distinctly conceive X existing in a certain way, then X can exist in that way, at least by God's power." This premise is similar in form to the statement "If I can see normally, then I can see molecules, at least with the help of a microscope." But for this statement to be true, microscopes do not have to exist: the statement was not *false* before microscopes came into existence, and it would remain true even if they were all destroyed. Likewise, Descartes's premise can be true whether or not God exists. The same goes for step 2 of his argument, since it follows from premise 1. Significantly, then, the conclusion drawn in step 6 ("I am really a different thing from my body") does not depend for its truth on God's existence either.

It might seem, however, that even if Descartes's premises do not depend on God's existence for their truth, one must at least admit that the conclusion drawn in step 7 ("I can exist without my body, at least by God's power") loses all interest if God does not exist. But since God *may* exist even if it has not been proved that he does exist and since Descartes's argument is only meant to show the *possibility* of disembodied survival anyway, even (7) is not deprived of all its significance by the failure of Descartes's philosophical theology. It must be admitted, however, that the subtlety of Descartes's argument makes it quite difficult to assess its exact significance. We shall return to this matter in the final section of this chapter. The point we wish to emphasize for now is just that the *soundness* of Descartes's argument does not depend on God's existence.

To reinforce this point in one last way, it may be useful to look at the matter differently. Suppose that the phrase “at least by God’s power” were simply deleted from Descartes’s entire argument. Then would not the argument be just as strong as before? If so, then this suggests that there is no essential connection between Descartes’s case for dualism and his philosophical theology. Rather, his argument for dualism is, as previously noted, essentially an argument from the independent conceivability of mind and matter.

Let us now turn to a second possible objection to Descartes’s argument. This is the objection made by Antoine Arnauld, in the fourth set of *Objections to the Meditations*. Arnauld was generally very sympathetic to the Cartesian philosophy, but he acutely criticized a number of Descartes’s arguments. (As we have seen, it was he who most clearly raised the problem of the circle.) In a clear and forceful way, he presented what is probably still the most penetrating objection to the proof of the real distinction.

Arnauld could not see that Descartes’s proof was any improvement over what we have called his preliminary, abbreviated version of the argument. As Arnauld put it:

But so far as I can see, the only result that follows from this is that I can obtain some knowledge of myself without knowledge of the body. But it is not yet transparently clear to me that this knowledge is complete and adequate, so as to enable me to be certain that I am not mistaken in excluding body from my essence. (CSM II 141, M 108, SPW 143, AT VII 201)

To clarify his point, Arnauld went on to give a parody of Descartes’s argument:

Suppose someone knows for certain that the angle in a semi-circle is a right angle, and hence that the triangle formed by this angle and the diameter of the circle is right-angled. In spite of this, he may doubt, or not yet have grasped for certain, that the square on the hypotenuse is equal to the squares on the other two sides; indeed he may even deny this if he is misled by some fallacy. But now, if he uses the same argument as that proposed by our illustrious author, he may appear to have confirmation of his false belief, as follows: . . . [H]e may say . . . “I know . . . that

everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it. And hence the fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct, since they are capable of being separated by God." Yet I clearly and distinctly understand that this triangle is right-angled, without understanding that the square on the hypotenuse is equal to the squares on the other sides. It follows on this reasoning that God, at least, could create a right-angled triangle with the square on its hypotenuse not equal to the squares on the other sides.

I do not see any possible reply here, except that the person in this example does not clearly and distinctly perceive that the triangle is right-angled. But how is my perception of the nature of my mind any clearer than his perception of the nature of the triangle? He is as certain that the triangle in the semi-circle has one right angle (which is the criterion of a right-angled triangle) as I am certain that I exist because I am thinking.

Now although the man in the example clearly and distinctly knows that the triangle is right-angled, he is wrong in thinking that the aforesaid relationship between the squares on the sides does not belong to the nature of the triangle. Similarly, although I clearly and distinctly know my nature to be something that thinks, may I, too, not perhaps be wrong in thinking that nothing else belongs to my nature apart from the fact that I am a thinking thing? Perhaps the fact that I am an extended thing may also belong to my nature. (CSM II 142–3, M 109–110, SPW 143–4, AT VII 201–203)

The logic of Arnauld's objection can be summarized as follows. Suppose that premise (5) of Descartes's proof is replaced with

(5') I can clearly and distinctly conceive a right triangle existing apart from the square on its hypotenuse being equal to the sum of the squares on its other two sides.

Then it follows from (5') and line (2) of Descartes's proof that:

(7') A right triangle can really exist without the square on its hypotenuse being equal to the sum of the squares on its other two sides.

Since (7') is absurdly false, and is validly deduced from (2) and (5'), either (2) or (5') must be false. But (5'), as the example of the angle in the semi-circle is supposed to show, is true. So (2) is false; therefore Descartes's proof is unsound. (Notice also that if (2) is false, then so is (1), since (2) follows from (1)—and also that Descartes's criterion of truth is false, since (1) follows from it. Moreover, neither (4), (6), nor (7) can be established; all that remains of the argument is (3) and (5).)

In his fourth set of *Replies*, Descartes wrestles with Arnauld's objection for several pages. Without trying to cover all of his points, let us focus only on the most instructive ones.

Descartes's first point is that the example of the right triangle is not parallel to the case of mind and body. Arnauld's example, he says,

differs in many respects from the case under discussion. First of all, though a triangle can perhaps be taken concretely as a substance having triangular shape, it is certain that the property of having the square on the hypotenuse equal to the squares on the other sides is not a substance. (CSM II 158, M 110, SPW 145, AT VII 224)

Descartes's point is simply that having the square on its hypotenuse equal to the squares on its other sides is a *property* of a right triangle (we might call it "the Pythagorean property," after Pythagoras, who demonstrated the Pythagorean Theorem that describes this property), not another *substance*. Thus, the most that Arnauld's example can show is that (2) is false when "X" designates a substance and "Y" designates a property. It cannot show that (2) is false when, as in Descartes's argument, "X" and "Y" both designate substances.

Although the example of the right triangle does differ from the case of mind and body in the way Descartes indicates, his point is a weak one: it does not get to the heart of Arnauld's objection; for, as Anthony Kenny incisively points out,

[I]t may be replied that being extended is a property, and Arnauld's argument shows that being able to conceive mind without this property does not show that mind in fact lacks it.⁴

⁴ Anthony Kenny, *Descartes: A Study of His Philosophy*, p. 94.

In other words: since extension is a property, and Descartes is trying to prove that a mind can exist without this property, he cannot afford to allow that (2) is false *even when "Y" designates only a property.*

Later, however, Descartes gives a better reply to Arnauld. He says:

It is true that the triangle is intelligible even though we do not think of the ratio which obtains between the square on the hypotenuse and the squares on the other sides; but it is not intelligible that this ratio should be *denied* of the triangle. In the case of the mind, by contrast, not only do we understand it to exist without the body, but, what is more, all the attributes which belong to the body can be denied of it. (CSM II 159, M 112, SPW 146–147, AT VII 227; emphasis mine)⁵

[T]here is no way in which the triangle can be distinctly understood if the ratio which obtains between the square on the hypotenuse and the squares on the other sides is said *not* to hold. (CSM II 159, M 111, SPW 146, AT VII 225; emphasis mine)

To analyze this reply, we can make three points:

1. The reply turns on Descartes's interpretation of the phrase, "clearly and distinctly conceive (understand, perceive) X apart from Y." Descartes explains in several places that this phrase does *not* mean just "clearly and distinctly conceive X without thinking of Y"; for he realizes that from one's ability to do that, it does not follow that X can really exist without Y. Thus, in first set of *Replies* he grants that one can clearly and distinctly conceive of motion without thinking of a

⁵ There is a rather similar passage in one of Descartes's letters:

There is a great difference between *abstraction* and *exclusion*. If I said simply that the idea which I have of my soul does not represent it to me as being dependent on the body and identified with it, this would be merely an abstraction, from which I could form only a negative argument which would be unsound. But I say that this idea represents it to me as a substance which can exist even though everything belonging to the body be excluded from it; from which I form a positive argument, and conclude that it can exist without the body. And this exclusion can be clearly seen in the nature of the soul, from the fact that one cannot think of a half of a thinking thing. (Letter of May 2, 1644 to Mesland, CSMK 236, AT IV 120)

body: one can form a clear and distinct idea of motion without thinking of a moving body. But, he implies, it does not follow from this that motion can exist “on its own,” without being the motion of some body (CSM II 86, AT VII 120–21). This is why, in the proof of the Real Distinction, “clearly and distinctly conceive X apart from Y” cannot mean “clearly and distinctly conceive of X without thinking of Y.” Rather, it means “clearly and distinctly conceive X existing while conceiving that Y does not exist.” (It is to suggest this reading that, in formulating the steps of his proof, we put the word “existing” after “X.”) As Descartes says in the sixth set of *Replies*:

I found that the distinction between . . . mind and body . . . is much greater than the distinction between things which are such that when we think of both of them we do not see how one can exist apart from the other (even though we may be able to understand one without thinking of the other). For example, we can understand the immeasurable greatness of God even though we do not attend to his justice; but if we attend to both, it is quite self-contradictory to suppose that he is immeasurably great but not just. (CSM II 298–9, AT VII 443)

2. Descartes could have argued that the phrase, “clearly and distinctly conceive X apart from Y,” cannot mean “clearly and distinctly conceive X existing while conceiving that Y does not exist” in Arnauld’s (5’), since “I can clearly and distinctly conceive a right triangle existing while conceiving that the property of having the square on its hypotenuse equal to the squares on its other two sides—the Pythagorean property—does not exist” makes little if any sense. Rather, in Arnauld’s (5’), the phrase must have some meaning other than the one it has in Descartes’s (2). Therefore, Arnauld’s example fails to refute (2), because (7’) cannot be deduced from (2) and (5’) if the key phrase means one thing in (2) and something else in (5’). But this would not have answered the point that Arnauld’s example does refute (2) when “Y” designates a property.
3. To meet this point, Descartes allows the key phrase to have a meaning that fits Arnauld’s example; namely, “clearly and distinctly conceive X existing while conceiving that X is not Y” (i.e., that X does not have property Y). This differs from Descartes’s own interpretation of the phrase, since it does treat Y as a property that X might or

might not have, rather than insisting that Y be a thing. But by interpreting the phrase in this way, Descartes is able to cut to the core of Arnauld's objection; for he can now show that the objection fails to prove that (2) is false even when "Y" designates a property. For Descartes's premise (5) now means

- (5a) I can clearly and distinctly conceive myself existing while conceiving that I am not extended (i.e., that extension is not a property of me).

Arnauld's (5'), on the other hand, now means

- (5a') I can clearly and distinctly conceive a right triangle existing while conceiving that the square on its hypotenuse is *not* equal to the sum of the squares on its other two sides.

However, Descartes points out, while (5a) is true, (5a') is false; for although I can think of a right triangle, and presumably even conceive it clearly and distinctly, without any thought of the ratio between the square on its hypotenuse and the squares on its sides, I cannot clearly and distinctly conceive a right triangle while conceiving that the square on its hypotenuse is *not* equal to the squares on its sides. As he puts it, "there is no way in which the triangle can be distinctly understood if the ratio which obtains between the squares on the hypotenuse and the squares on the other sides is said not to hold." Therefore, Arnauld's example fails to show that (2) is false even when "Y" designates a property.

Admittedly, this reply invites further questions. One would like to know exactly why it is possible to clearly and distinctly conceive a right triangle without thinking of the Pythagorean property but not possible to clearly and distinctly conceive a right triangle while "denying" that property of it. Descartes's idea seems to be that when "X is P" follows from one's conception of X, there is something worse about conceiving that X is *not* P than about just not thinking of P while conceiving X. He seems to be committed to the following criterion for clear and distinct conception:

S clearly and distinctly conceives X only if, for every property P such that "X is P" follows from S's conception of X, it is not the case that S can conceive that X is not P.

Obviously, it would be desirable to have an account of clear and distinct conception to support this criterion. But even in the absence of such an account, the criterion seems quite plausible. There is something to be said for the idea that while one can have a clear and distinct conception of X without being aware of all the properties that follow from one's conception of X, one fails to have a clear and distinct conception of X if one conceives it *not* to have a property that follows from one's conception of X. Furthermore, it will not do to object that for all Descartes knows, perhaps "I am extended" does follow from his conception of himself; for Descartes's careful examination of his idea of himself in *Meditation II* seems to show that this is not the case (whereas a careful examination of his conception of a right triangle would show that the Pythagorean property does follow from that conception).

We may conclude, therefore, that Arnauld's objection fails to refute Descartes's argument. Since that objection is probably the most acute one ever raised against Descartes's argument, we may accept the argument, at least provisionally. We shall have more to say about it in our overall assessment of Cartesian Dualism in section 6.

3. Descartes's Proof of the Material World

If one accepts Cartesian Dualism, then one faces a difficult question that has come to be known as "the mind-body problem": What is the *relationship* between mind and matter, specifically between the mind and body of a single person? Before discussing this issue, however, we need to consider Descartes's views about the material world. How does he finally overcome his doubt about the very existence of such a world? And what are material objects really like, according to him? These questions are addressed, respectively, in the present and in the next section. Then, in the last two sections, we shall be ready to consider the mind-body problem.

In the paragraph that directly follows his proof of the Real Distinction, Descartes finally overthrows his doubt concerning material things. He does so, as we might expect, by advancing a proof of their existence. Descartes begins the paragraph very cautiously, by noting that he has certain "faculties," namely, imagination and sensation, which are distinct from him in the same way as the modes of a physical thing would be from that thing, but which do *not* require the existence of any physical things. These

faculties require only his existence as a thinking substance, for they are merely modes or properties of a thinking substance, as is shown by the fact that they cannot be conceived to exist without a thinking substance to which they belong. Next Descartes notes that he also “recognizes” (i.e., has the clear and distinct conception of) certain other faculties or powers, such as motion and change of shape, that *would, if they really exist*, require the existence of an extended or material substance; for they clearly require extension—are modes of extension—and hence can really exist only if extended substance exists. The question is whether any extended substance does exist.

Descartes finally turns to this question about one-third of the way into the paragraph, where he says:

Now there is in me a passive faculty of sensory perception, that is, a faculty for receiving and recognizing the ideas of sensible objects; but I could not make use of it unless there was also an active faculty, either in me or in something else, which produced or brought about these ideas. (CSM II 55, SPW 115, AT VII 79)

Here Descartes’s terminology may get in the way of understanding the point he is making. By “ideas of sensible objects,” he means sensory experiences, which he here regards as a type or subclass of ideas, just as he did in *Meditation III*, where he classified “hearing a noise,” “seeing the sun,” and “feeling the fire” as falling among his adventitious ideas (CSM II 26, SPW 89, AT VII 38). Thus, as we shall see in a moment, he ascribes to these ideas the same notion of “objective reality” as he ascribes to all other ideas. By saying that he has “a passive faculty of sensory perception,” he means that he has the capacity to obtain these ideas of sensible objects without any awareness of producing them himself. Thus, his basic claim is just that he has ideas of sensible objects (= sensory experiences) that he is not aware of actively producing or conjuring up himself. Now, claims Descartes, these ideas must have some cause: his point after the semicolon is not merely the pragmatic one that if there were no cause of his sensory experiences, his “faculty” or capacity for “receiving and recognizing” the ideas would be *useless* to him, but, rather, that it would not even be *activated*, so that he would never even have the ideas. Descartes, then, is here relying on the principle that his ideas must have some cause—a principle that (as we have seen) is also central to his *Meditation III* proof of God’s existence.

We may now summarize the first step of Descartes's argument:

- (1) I have ideas of sensible objects that I do not seem to produce myself; these ideas must be produced by some cause.

The question now becomes: What is this cause? Descartes will argue that it must be material things. His argument proceeds by a process of elimination, that is, by ruling out all other possible causes of the ideas. We shall present it in a somewhat informal manner, because its formal structure is not as important as its overall conception.

The first possibility that Descartes seeks to rule out is that he himself is the cause of the ideas. (This possibility is not ruled out by (1), which only says that he does not *seem* to produce the ideas himself.) We may paraphrase what he says, at the start of the sentence immediately following the one quoted above, and incorporating the helpful emendation about "not presuppose[ing] any thought on my part" that Descartes substituted into the French version:

- (2) The cause of my ideas of sensible things is not in me as a purely thinking thing, because (a) it does not presuppose my thought, and (b) the experiences it produces come independently of my will.

To understand this, we need to know what Descartes means by something's "presupposing" thought. Something presupposes thought if it could not exist unless thought existed. For example, doubting, believing, and feeling all presuppose thought, because they could not exist unless thought existed. This is because they are specific kinds of thought or, in Descartes's own terminology, "modes" of thought (just as squareness and circularity presuppose extension, because they are specific forms, or modes, that extension can take on and that cannot really exist unless extension exists). What Descartes means by (a), then, is that whatever the cause of his ideas of sensible objects may be, it is not a mode of thought. For this cause is something that produces sensory experiences; it is an "active faculty." But to doubt, to believe, or to feel is manifestly not to produce or cause some sensory experience; and in (a) Descartes is making the general point that *no* mode of thought consists in causing sensory experiences. In point (b), he anticipates a possible objection to this generalization. The objection is that there is a mode of thought that

could be the cause of his ideas of sensible objects, namely, the will. In other words, perhaps I deliberately “will” (i.e., conjure up) these ideas. Descartes’s reply to this objection is that it is simply false that I will or conjure up the ideas, because they manifestly occur quite independently of my will.

It is important to understand that step 2 depends on the results of the proof of the real distinction; for Descartes is assuming that he *is* a purely thinking substance (or in any case that the part of him which thinks is distinct from any body that may also exist) and giving reasons why the cause of his ideas of sensible objects must be something other than this thinking substance. In other words, (2) has the quite limited function of showing that the cause of these ideas is not the thinking substance whose existence was asserted in the *cogito* and whose distinctness from anything physical was finally shown by the proof of the Real Distinction. Beyond this negative claim, (2) tells us nothing about the identity of this cause.

Descartes’s next step goes as follows:

- (3) The cause of my ideas of sensible things must therefore be some substance other than myself. Furthermore, this substance must contain, either formally or eminently, all the reality that the ideas it produces contain objectively.

Here Descartes uses once again the principle that the cause of an idea must have at least as much formal reality as the idea contains objective reality; for, as we saw in chapter 3, whenever *X* contains formally or eminently all the reality that idea *I* contains objectively, *X* has at least as much formal reality as *I* contains objective reality, and so it is possible for *X* to cause *I*.

Next, Descartes gives a breakdown of all the possible causes of his ideas of sensible things allowed by (3):

- (4) this substance is either

- (a) body (i.e., matter), a substance that contains formally all the reality that the ideas it produces contain objectively; or
- (b) God Himself, in which case it is a substance that contains eminently all the reality that the ideas it produces contain objectively; or

- (c) some created thing “more noble than body,” in which case, again, it is a substance which contains eminently all the reality that the ideas it produces contain objectively.

There seems to be a minor difficulty here. Descartes seems to overlook the possibility that the substance that causes his sensory experiences could be some created thing which is different from body but *exactly as noble* as (instead of “more noble” than) body. For example, suppose that this substance were some finite, nonphysical substance other than himself. Then Descartes seems to assume that this substance would have to possess more formal reality than the ideas it causes contain objective reality—to be “nobler” than the bodies these ideas portray. But the basis of this assumption is not obvious: Why couldn’t the substance in question possess exactly the same degree of formal reality as the ideas it causes contain objective reality? It would seem that in order to cover this possibility, Descartes should have formulated option (c) this way:

- (c') some created thing other than body, which contains eminently all the reality that the ideas it produces contain objectively.

(Remember that according to the explanation of “eminently contains” offered in chapter 3, a cause that contains exactly as much but a different kind of formal reality than its effect contains that effect’s reality “eminently.”)

Despite this somewhat esoteric difficulty, however, the general pattern of Descartes’s reasoning in step (4) is clear enough: his ideas of sensible objects could be caused by either (a) physical objects, (b) God himself, or (c) some other thing created by God—some “deputy” of God, so to speak.

In his next step, which is the crucial step in the argument, Descartes eliminates (b) and (c). We may paraphrase what he says as follows:

- (5) The cause of my ideas of sensible objects cannot be God or any created substance other than bodies; for God has given me no way to spot that this is so, but, instead, a very powerful inclination to believe that these ideas come from bodies (material objects). So God would be a deceiver if the ideas were produced in any other way. But since God is a supremely perfect being, he cannot be a deceiver.
- (6) Therefore, bodies exist.

If, as Descartes had supposed might be the case in *Meditation I*, all of his sensory experiences were caused by God himself or, so to speak, by some deputy of God, Descartes would have absolutely no way to detect this: he would be subject to a permanent, undetectable hallucination. Moreover, he would still have a virtually irresistible feeling—what the twentieth-century Spanish-American philosopher George Santayana called “animal faith”—that his sensory experiences did come from bodies. (Try actually to doubt that the sensory experiences you have at this very moment are caused by a written page. Is this not very difficult to do? This illustrates Descartes’s and Santayana’s point.) Thus, he would be irremediably deceived. But this means that God, who Descartes assumes has been clearly proved by his previous *Meditations* to be the creator of Descartes and whatever else exists, would be a deceiver; for a deceiving God would be precisely one who allows “any falsity in my opinions which cannot be corrected by some other faculty supplied by God” (CSM II 55–6, SPW 116, AT VII 80). But Descartes also takes himself to have abundantly shown that God is a perfect being, who therefore cannot be a deceiver, because, as he said in the *Third Meditation*, “It is manifest by the light of nature that all fraud and deception depend on some defect” (CSM II 35, SPW 98, AT VII 52). It follows, then, that sensory experiences are caused by material things: the doubt of the existence of material things generated in *Meditation I* is, at last, overthrown.

Having so proved the existence of the material world, Descartes immediately goes on to add an important qualification, which we may paraphrase as follows:

- (7) Material things may not be exactly as we perceive them to be by our senses, since such perception is often obscure and confused. But they must really contain all that is clearly and distinctly perceived in them by the mind, i.e., the geometrical properties—extension and its modes.

This remark sets the stage for Descartes’s views about the *nature* of material things, which will be considered in the next section.

The conclusion drawn in (6) is the very general one that “corporeal things” (bodies, physical things) exist. This conclusion does not tell us what physical things there are, or even whether there is only one physical thing in the universe or more than one. But Descartes now also feels entitled to draw certain more specific conclusions that, as he puts it, “I am

taught by nature." By this phrase, which must not be confused with "the light of nature" (= reason, or the faculty of clear and distinct perception), Descartes is indicating that there are many things which he is naturally or spontaneously inclined to believe. Now that he has overcome his generalized doubt about the existence of matter, he believes he can safely accept some of these things. It is important to note, before listing these, that Descartes says he can accept them just because of "the very fact that God is not a deceiver" (CSM II 55, SPW 116, AT VII 80). In other words, Descartes now takes himself to know for certain, solely on the grounds that God is not a deceiver, several specific things about the material world that he finds himself naturally impelled to believe. The things he lists are these:

1. I have a body.
2. I am very closely joined to this body.
3. There are other bodies (physical things).
4. My perceptions of colors, sounds, tastes, temperatures, and hardnesses enable me correctly to infer that the physical things that cause these perceptions have properties that vary as widely as, but may not resemble, those perceptions.
5. These bodies can affect me both beneficially and harmfully.

Points 1, 2, and 5 relate to Descartes's views about the relationship between mind and body. We shall say more about this important issue in section 5. Point 4 relates to Descartes's views about the nature of the material world, specifically, to a view called the theory of primary and secondary qualities. We shall discuss this topic in the next section. First, however, we need to assess Descartes's proof of the material world.

Given what Descartes has argued in his previous *Meditations*, his proof of the material world has a certain cogency. It makes a logical use of theses for which he has already argued, such as the existence of a perfect God and the real distinction between mind and body. It is a natural development and culmination of the overall argument of the *Meditations*.

This is not to say, however, that the proof is invulnerable, even within the framework of Descartes's own theism and dualism. One possible objection, for example, was made by the German Rationalist philosopher Leibniz (1646–1716). Leibniz questioned Descartes's claim that a perfect God would not allow us to be deceived about the causes of our sensory experiences. Perhaps, Leibniz suggested, God has certain good reasons

for allowing us to be so deceived, just as he has good reasons for allowing certain other evils. Perhaps the deception even works for our benefit.⁶

Even if Descartes could have answered this sort of criticism, there remains a much more fundamental objection. In the proof of the real distinction, as we have seen, God serves only to guarantee the truth of the clear and distinct perceptions contained in the proof. By contrast, the proof of the material world does not rely on God only as a guarantor of clear and distinct perceptions. Rather, it relies on God, in step 5, to rule out the skeptical hypothesis that despite our powerful inclination to believe that our sense experiences are caused by bodies, they may really be caused by an evil deceiver, or in some other bizarre way. So if there may be no God, then this skeptical hypothesis remains in place: no reason has been given that removes it. But the thesis that there is a God depends on the proofs of the existence of God in *Meditations III* and *V*, and those proofs are unsuccessful. Therefore, Descartes's proof of the material world must be regarded as a failure. The implication is that the radical doubt of *Meditation I* remains unanswered. For all Descartes knows, there may be no material world, and his sensory experiences may be produced by the deceiver or in some other extraordinarily bizarre way; for all he can be absolutely certain of, the only thing that exists is the purely thinking self of the *cogito*.

The significance of the failure of Descartes's proof of the material world for philosophy after Descartes can hardly be understated. Most major philosophers since Descartes have wrestled with the "problem of the external world" that he uncovered but failed to solve. Some, like Locke in the seventeenth century and Russell in the twentieth century, have tried to show that material things are the causes of our sense experiences by a quasi-scientific argument "to the best explanation" that does not appeal to God.⁷ Some, like Berkeley (1685–1753), John Stuart Mill (1806–1873) and A. J. Ayer (1910–1989), have argued that material things are in some way composed out of, or constructed from, the sense experiences themselves.⁸ Others, like David Hume (1711–1776), have

⁶ Gottfried Leibniz, "Critical Remarks Concerning the General Part of Descartes' Principles," p. 41.

⁷ For critical discussion of this approach, see Georges Dicker, *Berkeley's Idealism*, pp. 42–63 and pp. 194–201 and *Perceptual Knowledge*, pp. 95–118.

⁸ For critical discussion of this approach, see Dicker, *Berkeley's Idealism*, pp. 240–249 and pp. 273–80 and *Perceptual Knowledge*, pp. 123–55.

held that Descartes's problem is insoluble and that skepticism is the only rational position.⁹ Immanuel Kant (1724–1804) argued that the mere ability to order our own past experiences in time proves that those experiences are caused by successive states of physical objects that we perceive in space.¹⁰ Still others philosophers, like Dewey (1859–1952) and Wittgenstein (1889–1951), have argued that Descartes's doubt of the material world is illegitimate and should not have arisen in the first place.

4. Descartes on the Nature of the Material World

Although Descartes's proof of the existence of material things is a failure, it does not follow that his views about their nature are false or uninteresting. It is to these views that we turn in the present section.

As we saw in our discussion of the wax example of *Meditation II*, Descartes held that the only property included in a clear and distinct conception of body is extension, or three-dimensionality. For Descartes, bodies are, so to speak, units of extension, or units having three-dimensional shape and size. His conception of matter is thus a purely geometrical one. As he puts it in *The Principles of Philosophy*:

The nature of body consists not in weight, hardness, colour, or the like, but simply in extension. . . . [T]he nature of matter, or body considered in general, consists not in its being something which is hard or heavy or coloured, or which affects the senses in any way, but simply in its being something which is extended in length, breadth and depth. (CSM I 224, SPW 190, AT VIII A 42)

This conception of matter raises many questions, of which the most fundamental are probably these three:

⁹ For an analysis of Hume's skeptical position on the belief in an external world, see Georges Dicker, "Three Questions about *Treatise* 1.4.2," 115–53 and *Hume's Metaphysics and Epistemology*, pp. 154–89.

¹⁰ For a defense of this approach, see Georges Dicker, "Kant's Refutation of Idealism," 80–108, "Kant's Refutation of Idealism: a Reply to Chignell," 715–83, "Kant's Refutation of Idealism: Once More unto the Breach," 191–95, and *Kant's Theory of Knowledge*, pp. 194–211.

1. How do properties other than extension and its modes, such as color, taste, smell, sound, and heat and cold, “fit in?” Are they merely illusions of our senses? Or do they in some way also belong to bodies, and if so how?
2. How does matter differ from space, given that space, like Cartesian “body,” seems to be nothing but extension?
3. How does one body differ from another body, given that bodies are just units of extension? What demarcates one such “unit” from another?

In the three subsections to follow, we shall consider Descartes’s answers to each of these questions.

4.1 Primary and Secondary Qualities

In order to answer the first question, Descartes resorted to a theory that can be traced back to the Greek atomist Democritus (460–370 b.c.), that was accepted by nearly all seventeenth-century philosophers and scientists, and that is, in one form or another, still defended by many philosophers today—the theory of primary and secondary qualities. The classic exposition of this theory is given by John Locke (1632–1704) in his *Essay Concerning Human Understanding*, written some years after Descartes’s death.¹¹ In the following discussion we shall draw substantially on Locke’s account, in order to put Descartes’s version of the theory into better perspective. But since Locke’s account, despite its insightfulness, is ambiguous and sometimes even inconsistent, our exposition of it will be somewhat modernized and idealized, and we will take note of its principal inconsistency at the end of this subsection.

The theory makes a fundamental distinction between two kinds of qualities or properties of material things: primary qualities and secondary qualities. Primary qualities are defined as those that a thing *must* possess in order to be a *physical* object. According to Locke, these include shape, size, solidity, mobility, and number. Imagine, says Locke, that a grain of wheat is divided in half, that each half is again divided in half, and so on as far as you like. No matter how long this process of division continues, each part must still have some size, some shape, solidity, the ability to be moved; and, of course, there is always some number of parts.

¹¹ John Locke, *An Essay Concerning Human Understanding*, Book 2, chapter 8.

This shows, as Locke puts it, that these qualities are “inseparable from [a] body, in what estate soever it be.”¹² In other words, the primary qualities are those that are supposed to belong necessarily to any portion of matter whatsoever: they serve to *define* matter. As we might expect in view of Descartes’s “geometricized” concept of matter, his list of primary qualities is shorter than Locke’s. For Descartes, the primary qualities include only extension and its modes, plus motion. Notice, then, that for Descartes *solidity* is not one of the primary qualities.

The *list* of secondary qualities includes at least: colors, tastes, smells, sounds, heat and cold. Descartes adds solidity to the list. As we shall see, however, there are strong reasons for agreeing with Locke that solidity is a primary quality. So let us leave solidity out of our list of secondary qualities; we shall try to justify this stipulation in the next subsection. The classification of heat and cold as secondary qualities is also problematic because, at least today, we think of these as falling along the scalar magnitude that we call “temperature,” and arguably any body must have *some* temperature or other, in which case temperature is a primary quality and heat and cold are determinates of it. But Locke may have thought of heat and cold rather as opposites that characterize some but not all bodies, in which case they would qualify as secondary qualities. That said, we shall not pursue this point further, since it does not affect the basis contours of the theory.

The *definition* of secondary qualities, which is the most important feature of the entire theory, is that secondary qualities are only *capacities, powers, or dispositions* of physical objects to cause experiences of color, taste, smell, sound, and heat and cold in a perceiver, under normal conditions of observation. For example, the color red is only the capacity of certain objects (e.g., ripe tomatoes and fire engines) to cause experiences of red in a normal perceiver under normal light; and a sweet taste is only the capacity of certain objects (e.g., sugar cubes and ripe strawberries) to produce experiences of sweetness in a perceiver under normal conditions for gustatory perception. To introduce a piece of contemporary terminology, the definition of secondary qualities says that they are *dispositional properties*. A dispositional property (or, for short, a “disposition”) is a capacity to cause or to undergo some change. For example: *explosiveness* is a dispositional property of gunpowder, *fragility* is a dispositional property of glassware, *solubility* in water is a dispositional property of sugar,

¹² Locke, *Essay*, Book 2, chapter 8, section 9.

and *corrosiveness* is a dispositional property of some acids. The definition of secondary qualities, then, says that they are dispositional properties of a special kind: they are capacities to cause certain experiences in a sentient organism. Of course, in order for such capacities to operate, the right conditions must be present. For instance, in order for a red object's capacity to cause experiences of red to operate, there must be an appropriately situated perceiver equipped with normally functioning eyes, a sufficient amount of light, and so forth. But the same thing is true of other kinds of dispositional properties; for example, in order for the corrosiveness of an acid to operate, there must be a piece of metal of a certain sort, oxygen, a temperature not so cold that the acid freezes or so hot that the metal melts, and so forth.

As the terms “primary” and “secondary” may suggest, the theory also holds that the secondary qualities *depend* upon the primary qualities. In other words, it holds that whether or not an object has, say, the capacity to cause experiences of red in normal light depends upon whether or not it possesses certain primary qualities. Locke, who accepted the seventeenth-century atomism (“corpuscularism”) of scientists like Boyle and Galileo, ties this point to their atomic theory of matter. He says that the capacities in question depend upon the primary qualities of *the object's atomic parts*. As he puts it:

[S]uch qualities which in truth are nothing in the objects themselves but powers to produce various sensations in us by their *primary qualities*, i.e., by the bulk, figure, texture and motion of their insensible parts, as colours, sounds, tastes, etc. These I call *secondary qualities*.¹³

This contains two statements followed by Locke's definition of secondary qualities: (1) that physical things are composed of “insensible parts” (i.e., atoms or corpuscles), each of which have the primary qualities: a certain size, shape, solidity and velocity and (2) that because of the primary qualities of these atomic parts, physical things have “powers” (the word Locke always uses for capacities or dispositions) to cause perceivers to experience sensations of color, sound, taste, smell, heat and cold. It is these specific powers or capacities that Locke calls “secondary qualities.”

¹³ Locke, *Essay*, Book 2, chapter 8, section 9.

What is the rationale for the theory of primary and secondary qualities? There are at least two reasons that support the theory. The first is that it provides a way of fitting the secondary qualities into a scientific account of matter. Scientific descriptions of matter generally do not refer to secondary qualities. For example, no atomic theory of matter, whether ancient, seventeenth-century, or contemporary, ascribes such qualities as color, smell, or taste to the atoms themselves. These qualities are regarded by science as phenomena *to be explained* by the atomic structure of matter and must therefore not be ascribed to the atoms themselves. How, then, do these qualities relate to matter as it is described by physics? The answer proposed by Locke is a plausible one: colors, tastes, smells, and so on are really only capacities that objects have to affect perceivers in certain ways, because of their particular atomic structure. For example, suppose that a certain tabletop is brown. Then Locke would say that its being brown consists in its having the capacity to cause experiences of brown in us under normal conditions and that it has this capacity *because* of the molecular structure of its surface.

Notice that the scientific motivation for the theory of primary and secondary qualities need not necessarily be tied to an atomic theory of matter. We can see this by looking at Descartes's version of the theory. Unlike Locke, Descartes was not an atomist; for he believed that extension can be infinitely divided, whereas atoms would have to be indivisible particles. Nonetheless, scientific concerns are no less central to his version of the theory than to Locke's. For Descartes, the scientific description of matter is the clear and distinct one of matter as extension. Color, taste, smell, sound, and heat and cold are no part of this conception. So again, how do these qualities relate to matter as conceived by science? Descartes's answer, in his *Principles of Philosophy*, is that

[T]he properties in external objects to which we apply the terms light, colour, smell, taste, sound, heat and cold—as well as the other tactile qualities . . . —are, so far as we can see, simply various dispositions in those object which make them able to set up various kinds of motions in our nerves . . . which are required to produce all the various sensations in our soul. . . . [L]ight, colour, smell, taste, sound and tactile qualities . . . are nothing else in the objects—or at least we cannot apprehend them as being anything else—but certain dispositions depending on size, shape and motion. (CSM I 285, SPW 206, AT VIII A 322–23)

In other words, secondary qualities are merely dispositions of objects—which scientifically considered are composed only of shapes and sizes in motion—to affect perceivers in certain ways.

The scientific aspect of the theory of primary and secondary qualities explains some of the remarks Descartes makes after giving his proof of the physical world. As we saw in the previous section, immediately after the proof Descartes suggests that physical things are not exactly as we perceive them by our senses. A bit later he makes this more specific, saying that bodies have certain “differences” that explain, but do not resemble, the colors, tastes, sounds, smells, temperatures and hardnesses that we perceive. These points are implications of the theory of primary and secondary qualities. Descartes means that the experiences of color, taste, smell, and so on that bodies cause in us do not resemble the various shapes, sizes and motions in virtue of which bodies (have the dispositions to) cause such experiences in us. Essentially the same point was more prominently and famously made by Locke, who put it by saying that the “ideas” of secondary qualities do not resemble anything in material objects. Locke also claimed that, by contrast, the “ideas” of primary qualities do resemble those qualities themselves.

The second reason that supports the theory of primary and secondary qualities comes to light if one reflects on the following question: What is a color, such as the redness of an apple, when no one is looking at it? At such times, the apple is obviously not causing any experiences of red. To put this point more strikingly, the apple does not then *look* red; for there is no one *to whom* it looks red.¹⁴ What then is it for an apple to be red, when no one is looking at it? This question is much like the better-known puzzle, Does a tree that falls in a forest with no one there to hear it make a sound? A plausible answer to this puzzle is that if by “making a sound” one means having the capacity to cause an auditory experience in a perceiver (by producing vibrations in the air, called sound waves, which can stimulate the auditory sense receptors), then the falling tree does make a sound; for it does have this capacity (while it is actually breaking and crashing through the forest and into the ground.) But if by “making a sound” one means actually causing an auditory experience (or sounding in some way), then the tree makes no sound, because it cannot cause an auditory experience unless a perceiver is present. In other words, the tree does not then *sound* in some way, because it cannot sound in any way at

¹⁴ Compare Kenny, *Descartes*, p. 219.

all unless it sounds some way to someone. Thus the tree does make a sound; but this only means that it has the capacity or disposition to cause an auditory experience in a perceiver—that if a perceiver were present the tree would cause such an experience or would sound in some way to that perceiver. Now, the theory of primary and secondary qualities offers a parallel answer to our question about color (and to strictly analogous questions that could be asked about taste, smell, and heat and cold). A red object that no one is looking at is still red, because it still has the disposition or capacity to cause an experience of red in a normal perceiver under normal conditions; it is still true of the object that if a perceiver were to look at it under normal conditions, it would cause such an experience in him or her. But of course, a red object that no one is looking at doesn't then cause any such experience. It doesn't then *look* or *appear* red, since there is no one to whom it looks or appears red. Thus, if we ask, what exactly is a thing's redness, insofar as it exists on the thing's surface whether or not anyone is looking at it, the answer is that it is the power, capacity, or disposition of the thing to cause an experience of red in a perceiver (to look or appear red to a perceiver) under appropriate conditions. The thing has this disposition whether or not someone is looking at it; so it is red whether or not it is being perceived. And if we add that it has this disposition because of the molecular structure of its surface, then we have Locke's view that a thing's secondary qualities depend upon the primary qualities of its atomic parts.

Although the dispositional account of secondary qualities just given has much to recommend it, it needs supplementation. For as it stands it implies that colors, sounds, tastes, and smells are *only* powers, capacities or dispositions of a special sort, but that view is too narrow, for at least two reasons. One reason, which might be called the "liking-it" argument, is this. Suppose I tell you that my favorite color is aquamarine, or that I like the taste of chocolate. Surely I am not then saying that I like the power, capacity or disposition of things to produce certain visual or gustatory sensations. The mere capacity or disposition is not what I care about, like, or enjoy. Nor does what I say mean simply that I like *things that have* these powers or dispositions. For aquamarine could be my favorite color even though I disliked aquamarine things because (say) they all came in ugly shapes, and I could like the taste of chocolate even though I disliked things that had that taste because they all contained an unpalatable stuffing. But if, when I say that aquamarine is my favorite color or that I like the taste of chocolate, I am talking neither

about certain dispositions nor about the things that have those dispositions, then what am I talking about? The answer seems to be that I am talking about a distinctive, qualitative aspect of aquamarine things and of chocolate—an aspect that I cannot define verbally but that I am familiar with from experience, and that cannot be identified with or reduced to a mere power or capacity.

The other reason why the dispositional account of secondary qualities is too narrow comes out if one reflects on the following passage from Berkeley's *Principles of Human Knowledge*:

[C]an there be a nicer strain of abstraction than to distinguish the existence of sensible objects from their being perceived, so as to conceive them existing unperceived? Light and colours, heat and cold, extension and figures, in a word, the things we see and feel, what are they but so many sensations, notions, ideas or impressions on the sense; and is it possible to separate, even in thought, any of these from perception? For my part, I might as easily divide a thing from itself.¹⁵

One way to read this passage is as a challenge, which we may call Berkeley's Challenge, to distinguish the existence of any sensible quality (what Berkeley called its *esse*) from its being actually perceived (what he called its *percipi*).¹⁶ Berkeley's challenge says, in effect: "you cannot even conceive a color existing apart from its being seen, a sound existing apart from its being heard, an odor existing apart from its being smelled, etc." Berkeley uses his challenge to support his own idealist view that all the qualities we perceive are really only ideas or sensations in our minds.

Berkeley's challenge seems ineffective when applied to primary qualities—why can't one conceive, for example, of an unperceived cube simply as a shape that fills a certain volume, or of an unperceived solid object as one that excludes other bodies from the space it occupies? But the puzzle about the tree that falls in a forest with no one to hear it shows that the challenge has at least *some* force with respect to sounds, and analogous puzzle cases can be described for tastes, smells, and even colors. To see this better, consider the following thought experiment.

¹⁵ George Berkeley, *A Treatise Concerning the Principles of Human Knowledge*, p. 104.

¹⁶ For a defense of this way of reading the passage, see Dicker, *Berkeley's Idealism*, pp. 75–82.

Suppose that there existed only one red thing in the entire universe, that a normal perceiver saw that thing in sunlight at a time t_1 , but that no one at all saw it at time t_2 . Then there is a sense in which the quality, red(ness), would exist at t_1 but not at t_2 ; likewise there is a sense in which tastes, sounds, and smells would vanish from the universe if (or at least during the times when) no one perceived any objects having those qualities. We suggest that one test that any adequate view about the nature of colors, tastes, sounds and smells should meet is to make clear in *what* sense this is true. The purely dispositional account of them that we have so far given does not meet that test, since objects continue have the dispositions to cause the relevant experiences even when they are not being perceived.

In order to meet the test, we should recognize that colors, tastes, smells, and sounds have a certain complexity; they have two aspects. On the one hand, they are capacities, powers, or dispositions of things to appear in certain ways to sentient beings under certain conditions—this we may call their *dispositional aspect*. (Note that this aspect corresponds to Locke's definition of a secondary quality.) On the other hand, they are the manifestations of those dispositions—this we may call their *manifest aspect*. For example, the color red is the disposition of some objects, such as fire engines, traffic stoplights, and ripe tomatoes, to look red to normal perceivers in standard light. But red(ness) is also the distinctive qualitative event or episode, involving consciousness, that occurs when the conditions are right for this disposition to manifest itself, that is, when there is a perceiver with normal vision looking at a red object in normal light. Accordingly, there are *two* true answers to the question: “What is the color red?” One answer is that red is the disposition of certain objects to look red to normal perceivers in normal light. The other answer is that red is the event, occurrence or episode that constitutes the manifestation of this disposition. We have chosen to speak of two “aspects” simply for lack of a better term. This term should not obscure the fact that the two aspects are related in a specific, unmysterious manner: the manifest aspect is the manifestation of the dispositional aspect, just as the process of corrosion is the manifestation of corrosiveness.

The distinction between the dispositional and the manifest aspect answers both the liking-it argument and Berkeley's Challenge. For what I like or enjoy if I like (say) red is the manifest aspect of red, not its dispositional aspect. Furthermore, what ceases to exist when no one perceives red is again its manifest aspect, not its dispositional aspect. But since the manifest aspect of red is in one sense identical with the quality red, it

provides a sense in which, just as Berkeley held, red(ness) exists only by being perceived. On the other hand, the distinction does not concede to Berkeley his idealist view that red, as well as any other sensible quality, is nothing but an idea; it preserves the objectivity of sensible qualities. For their dispositional aspect continues to exist in material things whether those things are perceived or not. Furthermore, even the manifest aspect of red is not purely subjective, since it exists only when an object's capacity to look red is manifested or exercised—that is, only when an object like a ripe tomato or a fire engine actually looks red to someone. Thus, an object's looking red can be categorized as a relational property of the object, that is, as a property that it can have only by standing in some specific relation to another object, akin to the property of being a sibling, being North of, or being taller than. In this case, the specific relation involved is that of looking red to someone; in other cases it might be that of smelling sweet to someone, or tasting bitter to someone, and so on.¹⁷

Locke was sensitive to the complexity of secondary qualities. But for reasons that would take us too far afield to pursue here, he does not make the distinction between their two aspects. Instead, he defines secondary qualities simply as powers or dispositions to cause sensations or ideas of color, taste, smell, sound, and heat or cold in us, and he sticks consistently to that definition. But his answer to the question “What is a color, a taste, a smell, a sound, or an instance of heat or cold?” is inconsistent. He oscillates between equating colors, tastes, sounds, smells, heat and cold with *only* the dispositions to cause in us sensations or ideas of colors, tastes, sounds, smells, heat and cold (i.e., only with what his definition calls “secondary qualities,” or what we have called the dispositional aspect of a secondary quality), and equating them with *only* those ideas themselves (i.e., only with what we have called the manifest aspect of a secondary quality, except that he construes the manifest aspect as an idea caused in a certain way rather than as a relational property of the object). Sometimes Descartes too is read as identifying colors, tastes, sounds, smells, and heat/cold with mere ideas, as did Galileo and other scientists of the day, though it is debatable whether this is a completely fair reading of him.¹⁸

¹⁷ See the next note.

¹⁸ There is a fuller account of primary and secondary qualities, of Locke's treatment of them, and of the distinction between the dispositional aspect and the manifest aspect of secondary qualities, in Dicker, *Berkeley's Idealism*, chapter 1. See also Georges Dicker, “Primary and Secondary Qualities,” 457–71, reprinted in Walter E. Creery, *George Berkeley*, pp. 27–43.

4.2 Matter, Space, and Solidity

We now turn to the second question raised by Descartes's "geometrized" concept of matter, namely, the question of how matter differs from space. This question is bound to cause trouble for Descartes; for how does body, if it is a substance whose whole nature is extension or three-dimensionality, differ from space, whose nature seems to be exactly the same? By defining matter as mere extension, has not Descartes abolished the difference between matter and space?

Descartes attempts to explain the difference between the two in his *Principles of Philosophy*, where he writes:

There is no real distinction between space . . . and the corporeal substance contained in it; the only difference lies in the way in which we are accustomed to conceive of them. For in reality the extension in length, breadth and depth which constitutes a space is exactly the same as that which constitutes a body. The difference arises as follows: in the case of a body, we regard the extension as something particular, and thus think of it as changing whenever there is a new body; but in the case of a space, we attribute to the extension only a generic unity, so that when a new body comes to occupy the space, the extension of the space is reckoned not to change but to remain one and the same, so long as it retains the same size and shape and keeps the same position relative to certain external bodies which we use to determine the space in question . . . [I]f a stone is removed from the space or place where it is, we think that its extension has also been removed from that place, since we regard the extension as something particular and inseparable from the stone. But at the same time we think that the extension of the place where the stone used to be remains, and is the same as before. (CSM I 227–8, SPW 193–4, AT VIII A 45–46)

Descartes is here making two points: (1) the difference between a body and the space or place it occupies lies only in the way we conceive of the two; (2) the body is that which we conceive as removable from the space, while the space is conceived as that which must remain when the body is removed from it. The first point already seems wrong, because it denies that there is a real difference (one independent of how we conceive things) between matter and space. But let us focus on the second point.

The key idea in it is that of removing something, say a stone, from the space or place that it occupies. If that idea fails to make sense, then Descartes's attempt to preserve even a merely "conceptual" distinction between space and body collapses. But the trouble is precisely that this idea, considered in light of Descartes's identification of both body and space with mere extension, does not make sense. For suppose we ask: When we conceive that the stone is removed from the place it occupied, exactly what, according to Descartes, do we conceive as being removed from the place? Descartes cannot answer that we conceive that something colored, or smelly, or noisy, or tasty, or hot or cold, is being removed from the place; for these secondary qualities are only capacities or dispositions grounded in the extension of bodies, and so we can understand their "removal" only if we can make sense of the removal of the extension itself—which is precisely our difficulty. Nor can Descartes say that we conceive that something solid is being removed from the place, because according to him, solidity also is merely a capacity, grounded in extension, to cause tactile sensations in a perceiver. The upshot is that we seem to be removing only a *volume* from the place—or a volume from a volume. But this seems quite unintelligible.

It would seem, then, that some property or properties other than extension must also be included in any adequate conception of matter. One property that naturally suggests itself is *solidity*, and this is indeed what Locke proposed:

[S]olidity . . . , of all others, seems to be the idea most intimately connected with and essential to body, so as nowhere else to be found or imagined, but only in matter. . . . By this *idea* of solidity is the extension of body distinguished from the extension of space: the extension of body being nothing but the cohesion or continuity of solid, separable, moveable parts; and the extension of space, the continuity of unsolid, inseparable, and immoveable parts.¹⁹

But, you may ask, what exactly is solidity? Well, one step toward clarifying this notion is to distinguish it from hardness. A hard body is one whose parts are not easily moved relative to each other; a soft body, like a pillow or a quantity of water, is one whose parts are easily displaced

¹⁹ Locke, *Essay*, book 2, chapter 4, sections 1 and 5.

relative to each other. Solidity, by contrast, consists in filling up space or in excluding all other bodies from the space actually occupied. It is not that a solid mass cannot be moved from the place it occupies, but rather that no other mass can occupy that same place so long as the first one still occupies it. So, even a soft body is solid. As Locke put it:

All the bodies in the world, pressing a drop of water on all sides, will never be able to overcome the resistance which it will make, as soft as it is, to their approaching one another, till it be moved out of their way: whereby our *idea* of *solidity* is distinguished both from *pure space*, which is capable neither of resistance nor motion, and from the ordinary *idea* of *hardness*.²⁰

Solidity, then, is virtually the same as impenetrability; whereas hardness is immovability relative to the immediately surrounding matter. To quote once more from Locke's insightful discussion:

Solidity is hereby also differenced from *hardness*, in that solidity consists in repletion, and so an utter exclusion of other bodies out of the space it possesses: but hardness, in a firm cohesion of the parts of matter making up masses of a sensible bulk, so that the whole body does not easily change its figure.²¹

Insofar as Locke is offering a criterion for distinguishing body or matter from space, his view seems plausible. But we must not rashly conclude that he has provided an adequate definition of "body" or of matter, for to say that something is solid seems to mean just that it excludes other bodies from the space it occupies. Thus if Locke is read as offering a definition, then his definition seems to define body in terms of solidity and solidity in terms of body, which is circular. Possibly this circularity could be avoided by defining solidity in a slightly different way, as the property of excluding all but one extended thing (including a portion of empty space) from a place. Then a body could be defined as an extended and solid thing. Extension would still have to be included in the definition since, as

²⁰ Locke, *Essay*, section 5.

²¹ Locke, *Essay*, section 4. Descartes sometimes says that the concept of body involves its excluding other bodies from the space it occupies (e.g., CSM II 17, SPW 81, AT VII 26). The question, however, is whether he is *entitled* to say this, given his purely geometrical account of the nature of matter.

Descartes and Locke both thought, a thing could not be a body without being extended. But the definition would not be circular since solidity would not be defined merely as the ability to exclude a *body* from a place, and a body could be thought of as a an extended thing that excludes all but one extended item (namely, itself) from the place that it occupies.

4.3 Bodies as Substances versus Bodies as Modes of Substance

Descartes's inability to differentiate matter from space leads naturally to our final question about his theory of the material world: how does one body differ from another, given Descartes's view of matter as extension? Descartes seems to have given two incompatible answers to this question. One answer, which he gives in his *Principles of Philosophy*, is the traditional one that one body differs from another because they are two different substances. The other answer, which he implicitly gives in the *Synopsis* of the *Meditations*, is the surprising one, which we already described in chapter 2, that two or more bodies are just different *properties* of a *single* substance. We shall examine these opposed passages and argue that only the latter answer is compatible with Descartes's doctrine that matter is merely extension.

In his *Principles*, Descartes defines a *real distinction* as one between two or more substances. (This is why he calls his proof that mind and body are two different substances a proof of the Real Distinction between them.) He then states his principle that “we can perceive that two substances are really distinct simply from the fact that we can clearly and distinctly understand one apart from the other,” from which he says it follows that “each and every part of [extended or corporeal substance], as delimited by us in our thought, is really distinct from the other parts of the same substance” (CSM I 213, SPW 180, AT VIII A 28). Descartes’s language is not careful here. In particular, the phrase “really distinct from the other parts of the same substance” is poorly chosen, since he has just defined a real distinction as one between *different substances*, not between parts of the same substance. Nevertheless, it seems evident that Descartes here means that any portion of extension that we can clearly and distinctly conceive apart from other portions is really distinct from—and hence a different substance from—other substances of the same kind (from the other “parts” of extension taken as a generic totality). This interpretation is confirmed in the next sentence, which reads, in part: “Each of us . . . is really distinct from every other thinking

substance and from every corporeal substance"; for the phrase "every corporeal substance" implies that there is a plurality of corporeal substances, each of which can only be a different body. This reading is further confirmed in the next principle, where Descartes says:

A different case . . . is the distinction by which the mode of one substance is different from another substance. . . . An example of this is the way in which the motion of one body is distinct from another body. . . . It seems . . . appropriate to call this kind of distinction a real distinction, since the modes in question cannot be clearly understood apart from the really distinct substances of which they are modes. (CSM I 214, SPW 181, AT VIII A 30)

Here Descartes gives, as an example of a case where the mode of one substance is distinct from *another substance* (i.e., other than the one to which that mode belongs), the case where the motion of one body is distinct from *another body* (i.e., other than the one to which that motion belongs). This implies that different bodies are different substances. The passage also implies that different bodies are "really distinct substances," which entails (both because "substances" is plural and because a real distinction is by definition one between two or more substances) that different bodies are different substances.

So far, we have merely made the textual point that Descartes says that different bodies are different substances. The question, however, is whether he is entitled to say this, given his view that bodies are nothing but extension. The answer, we suggest, is that he is not. To see why, note first of all that saying that different bodies are different extended substances obviously assumes that there can be more than one extended substance—that there can be a plurality of extended substances. We may safely assert, then,

- (1) If there cannot be more than one purely extended substance (i.e., substance consisting of nothing but extension), then the difference between two or more bodies cannot be that they are different purely extended substances.

But the difficulty is precisely that Descartes is not entitled to say that there can be more than one purely extended substance; for although extension can be divided into regions, it does not have separable parts. It makes no sense to talk of separating one part of pure extension from

another anymore than of separating one part of space from another: Where would you put it? We may also assert, then,

(2) Extension does not have separable parts.

But this implies that there could not be a plurality of purely extended substances, for what would make them different? We cannot say that just being different regions or portions of extension would make them different, because even a single extended substance must have different spatial portions—otherwise it wouldn't be extended. We cannot appeal to properties other than extension, like color or hardness. The only thing that could make two purely extended substances different would be that they could, at least in principle, be separated or disjoined from each other. But this is precisely what we cannot say about portions of pure extension. It seems clear, then, that we can say,

(3) If extension does not have separable parts, then there cannot be more than one purely extended substance.

There follows from (2) and (3), however,

(4) There cannot be more than one purely extended substance.

Further, there follows from (1) and (4),

(5) The difference between two or more bodies cannot be that they are different purely extended substances.

What, then, is the difference between two or more bodies? The only answer possible for Descartes seems to be that the bodies are different *modes* or accidental properties of a single, all-encompassing extended thing or substance (*res extensa*). This single extended substance, which might be called “matter-space” or “space-matter,” constitutes the entire physical universe. There may be many thinking substances or minds—as many as there are beings who could say or think “Cogito, ergo sum”—but there can be only one extended substance.

This “one-substance” view of the physical world is, interestingly enough, implicit in the passage from the *Synopsis* that we mentioned. In the context of explaining what would have to be done in order to prove that the soul is immortal, Descartes writes:

[T]he premises which lead to the conclusion that the soul is immortal depend on an account of the whole of physics. This is required for two reasons. First, we need to know that absolutely all substances . . . are by their nature incorruptible . . . [S]econdly, we need to recognize that body, taken in the general sense, is a substance, so that it too never perishes. But the human body, in so far as it differs from other bodies, is simply made up of a certain configuration of limbs and other accidents of this sort; whereas the human mind is not made up of any accidents in this way, but is a pure substance. For even if all the accidents of the mind change, so that it has different objects of the understanding and different desires and sensations, it does not on that account become a different mind; whereas the human body loses its identity merely as a result of a change in the shape of some of its parts. And it follows from this that while the body can very easily perish, the mind is immortal by its very nature. (CSM II 10, SPW 74, AT VII 13–4)

We have met (part of) this passage before; it is the place where Descartes uses the Argument from Change to show that the mind is a substance. And we shall meet the passage again in Section 6, when we offer an overall assessment of Cartesian Dualism. But what interests us now is what Descartes here says about body. He draws a major contrast between “body, taken in the general sense” and “the human body.” The former refers to the *totality* of extension. Descartes here regards this totality as a single, unitary, all-encompassing extended substance, or *res extensa*. He says that a complete physics could show that this substance, as well as the soul, is naturally incorruptible and therefore “never perishes.” By contrast, he says that a human body is “made up of *accidents*,” so that it can easily perish when some of those accidents are changed. In other words, Descartes here regards a human body as merely an aggregate of accidental properties or modes, which perishes when some of these modes are altered. Now there is every reason to suppose that Descartes would treat other sorts of bodies in the same way as human bodies. But in that case, Descartes is here implying that all particular bodies (e.g., chairs, rocks, planets, etc., as well as human bodies) are (*clusters of*) accidental properties, or modes. And if we ask what they are modes of, then the answer is that they are modes of the one, all-encompassing, incorruptible extended substance. The view that emerges is that there is only one extended substance and that particular physical objects are modes of it.

On this “one-substance” view of matter, the famous illustration of the wax in *Meditation II* would have to be interpreted differently than it was in chapter 2, section 4. Appearances to the contrary notwithstanding, the illustration would have to be seen as not concerned with what is required for a piece of wax to continue existing as the same wax or even for a chunk of matter to continue existing as the same chunk of matter. Rather, the piece of wax would have to be seen, as suggested in chapter 2, section 6.1, as a miniature model that Descartes uses to represent the whole physical universe. The point of the illustration would then be that as long as the physical universe retains the determinable property of extension, it continues to exist as the same physical universe. The underlying substance that remains the same throughout a process of change would be identical with the determinable property of extension. More strictly speaking (since the term “property” is often used to stand for the universal, which is then distinguished from its particular instances, exemplifications, or occurrences), the underlying substance would be identical with an instance of the determinable property of extension—the sole or unique instance, given the one-substance view. Since a determinate property cannot occur (cannot be exemplified or instantiated) unless its corresponding determinable property occurs, the principle that there are no “free-floating” properties (i.e., that if there is a property, then there must be a substance to which it belongs) would still hold, although it would now mean that if a determinate property occurs, then the appropriate determinable property must occur. Conversely, since a determinable property cannot occur unless one of its determinates occurs, the principle that a substance cannot exist without having some properties would also hold, although it would now mean that the determinable properties of extension and thought cannot occur unless there exist determinate shapes and sizes and determinate thoughts.

This way of interpreting Descartes receives some textual support from his *Principles of Philosophy*, part 1 (especially principles 53 and 63). In those places, Descartes seems to equate a substance with its defining attribute. Thus, in principle 53 he says that “each substance has one principal property which constitutes its nature or essence” (CSM I 210, SPW 177, AT VIIIA 25). And in principle 63 he says,

Thought and extension can be regarded as constituting the natures of intelligent substance and corporeal substance; they must then be considered as nothing else but thinking substance

itself and extended substance itself—that is, as mind and body.
(CSM I 215, SPW 182, AT VIIIA 30–1)

Here Descartes seems to be using the terms “extended substance,” “thinking substance,” “mind,” and “body” to stand partly for types, rather than for individual substances, minds, or bodies; and to be saying that these types are identical with their defining attributes. The passage also strongly suggests, however, that an *individual mind* is identical with an instance of the attribute of thought, and that an *individual body* is identical with an instance of the attribute of extension.

As noted earlier, on the view that a substance is identical with (an instance of) its defining attribute, Descartes’s version of the substance theory would not be vulnerable to the standard empiricist objection to substance that we discussed, namely that substance is in principle unperceivable.²² It might be argued that this objection would still apply, because a determinable property like extension cannot be perceived “by itself,” that is, without perceiving some determinate of it like a square shape or a circular shape. But even if it is true that a determinable cannot be perceived without perceiving one of its determinates, this does not mean that the determinable property is unperceivable. It means, rather, that to perceive a determinate property is also to perceive the corresponding determinable—for example, to perceive red is also to perceive color, to perceive squareness is also to perceive shape.

On the other hand, if (a) physical substance is identical with (an instance of) the determinable property of extension, then the substance theory can no longer even purport to give an account of the continuing existence of an ordinary physical object like a piece of wax or a stone. For, as we argued in chapter 2, section 6.1, retaining some shape or other, or some size or other, or some color or other, would not be enough for an ordinary physical object, for example your car, to continue existing even though all of its determinate properties had changed. But since on the view in question there is really only one immense physical object, the physical world itself, of which pieces of wax, stones, human bodies, cars, and so forth are only modes or accidental properties, the need to provide an account of such things’ identity through time is less pressing. Of course, a philosopher who wished to give an account of the identity through time of ordinary physical objects (despite their “modal” status)

²² See chapter 3, section 6.1.

could still do so, in terms of some feature other than substance (e.g., spatio-temporal continuity under a sortal, as also mentioned in chapter 2). But Descartes himself seems, in the end, not to have been very interested in this question.

It is noteworthy that Spinoza (1632–1677), a close student of Descartes's philosophy who felt that Descartes had failed to follow out consistently the logical implications of his own thought, proposed just such a one-substance view of the physical world as we have described. Spinoza, however, went even further. In his major work, entitled *Ethics Demonstrated in Geometrical Order*, he argued that the one extended substance is also the one and only thinking substance, which he called "God or Nature." The resulting view is that the entire universe consists of only one substance that is both thinking and extended.

To summarize: Descartes apparently held two incompatible views about bodies. One view, which is the traditional view derived from Aristotle and the Scholastics, is that each body is a distinct substance. The other view, which foreshadows Spinoza's theory, is that bodies are modes of a single substance. Only the latter view seems consistent with Descartes's doctrine that matter is merely extension. Finally, it should be noted that the former view becomes once again defensible if, as we suggested in the previous subsection, solidity is included in the conception of matter; for unlike a mere portion of space, one solid body can be separated from another, even if the two happen to be in contact.

5. Dualism and the Problem of Interaction

Descartes's dualism of mind and matter implies a certain conception of human beings or persons. A person is a composite entity, made up of two distinct components: a mind, or soul, and a body. Of course, Descartes gives a certain "priority," so to speak, to the mind; for he holds that he could exist without his body, from which he concludes (with the help of the principle that nothing without which a thing can exist is included in its essence) that his body is not part of his essence. By contrast, he would certainly deny that he could exist without his mind, and he insists that his mind *is* his essence. And in the *Synopsis* of the *Meditations*, as we saw in the last section, he goes so far as to say that while the mind is a substance, the body is composed only of "accidents." But despite Descartes's emphasis on the mind, he recognized that it would be fantastic to deny

that human beings are, at least during their earthly lives, embodied and that their embodiment is a salient fact of their existence. So, even while maintaining his sharp dualism, he tried to do justice to the close and intimate relation that each of us bears to his or her own body. This comes out, for example, in the fact that Descartes lists “I have a body” as the first and most obvious particular fact that he can accept about the material world and “I am very closely joined to this body” as the second. It also comes out in the language that he goes on to use in order to describe this close union:

Nature also teaches me, by these sensations of pain, hunger, thirst and so on, that I am not merely present in my body as a pilot in his ship, but that I am very closely joined and, as it were, intermingled with it, so that I and the body form a unit. If this were not so, I, who am nothing but a thinking thing, would not feel pain when the body was hurt, but would perceive the damage purely by the intellect, just as a sailor perceives by sight if anything in his ship is broken. Similarly, when the body needed food or drink, I should have an explicit understanding of the fact, instead of having confused sensations of hunger and thirst. For these sensations of hunger, thirst, pain and so on are nothing but confused modes of thinking which arise from the union and, as it were, intermingling of the mind with the body. (CSM II 56, SPW 116, AT VII 81)

Here and in some other places, Descartes goes for far as to imply that a purely thinking thing could not have sensations of pain, hunger, or thirst; rather, such sensations belong to a hybrid category that emerges from the union of mind and body. John Cottingham has emphasized this aspect of Descartes’s thought, calling it “Cartesian Trialism.”²³ The basic idea, which Cottingham explains lucidly and in detail, but which he also concedes does not fit well with Descartes’s radical dualism, is that although there are only two kinds of substance—mental and physical—some of the attributes of substance, especially sensation, fall into a third category that is neither purely mental nor purely physical but somehow emerges from a “substantial union” of mind and body.²⁴

²³ John Cottingham, *Descartes*, pp. 127–32.

²⁴ For a stimulating discussion of this kind of view, see Alison Simmons, “Re-Humanizing Descartes,” 52–71.

Descartes's desire to do justice to the fact of embodiment raises one of the most perplexing problems of his entire philosophy: What exactly is this relationship of substantial union between the mind and the body of a person? To appreciate why this question is such a difficult one for Descartes, remember that according to him the mind and the body are radically different and even opposite: mind is a thinking and nonextended substance, body is an extended and nonthinking substance. How then can there be a close "union" or "intermingling" of the two? What can this union possibly consist in? John Cottingham puts the difficulty this way:

The notion that two different substances can unite to form a single thing is not, in itself, obscure or problematic. We are familiar nowadays, for example, with the idea that hydrogen and oxygen can unite to form water; furthermore, this 'substantial union' generates 'emergent' properties—water has properties such as that of being drinkable which were not present in its constituent elements—and this (though it is not of course Descartes's own example) might be thought to give some grip on the Cartesian notion that events like sensations emerge or 'arise' when mind and body are united, even though they are not part of the essence of either *res cogitans* or *res extensa*. Nevertheless, Descartes himself clearly felt that his notion of the 'substantial union' of mind and body presented problems. For mind and body, as defined throughout [Descartes's] writings, are not just different, but utterly incompatible substances: in terms of their characteristics, they mutually exclude one another, since mind is defined as nonextended and indivisible, whereas matter is by its nature extended and divisible. And it is not easy to see how incompatible items can be, in any intelligible sense, 'united'. As Descartes rather ruefully put it:

it does not seem to me that the human mind is capable of conceiving at the same time the distinction and the union between body and soul, because for this it is necessary to conceive them as a single thing and at the same time to conceive them as two things, and this is absurd.”²⁵

²⁵ John Cottingham, "Descartes: Metaphysics and the Philosophy of Mind," p. 222. The quotation from Descartes is from his letter of June 28, 1643 to Princess Elizabeth (CSMK 227, AT III 692).

Moreover, Descartes's account of the mind-body union is ambiguous and at times even inconsistent. The chief ambiguity concerns whether the mind is joined to the whole body or only to a certain part of it. There are texts supporting each interpretation. In some places, such as the passage just cited, Descartes seems to hold that the mind is "intermingled" with the entire body; in other places he says that it is joined to the body only by virtue of being united with a particular part of the brain (compare, e.g., paragraphs 19 and 20 of the *Sixth Meditation*). Although strongly influenced by Descartes's overall position, later dualists like C. D. Broad (1887–1972) and C. J. Ducasse (1881–1969) have invariably assumed that only the latter view is plausible; indeed the former view is seldom, if ever, mentioned in recent presentations of dualism. No doubt this is partly because it seems quite impossible to take Descartes's talk of the "intermingling" of mind and body literally; for an nonextended mind cannot be intermingled with an extended body, as, for example, powdered milk can be intermingled with water: the mind would have to be composed of small *extended* parts that could be dispersed throughout the body. So it seems that Descartes's talk of intermingling can at best be taken as a way of describing how we *seem* to ourselves to be joined to our bodies, rather than as a literal, true account of the connection. Indeed, in the passage from *Meditation VI* just quoted, Descartes himself hedges his statements, saying that the mind is "as it were" intermingled with the body. It must be admitted, nevertheless, that sometimes Descartes does straightforwardly assert that the mind is joined to the whole body, as in the following passage from *The Passions of the Soul*, a late work in which he tried to describe the union of mind and body in some detail.

[T]he soul is really joined to the whole body, and . . . we cannot properly say that it exists in any one part of the body to the exclusion of the others. (CSM I 339, SPW 229, AT XI 351)

But whether the mind is joined to the whole body or only to some part of the brain, a much more fundamental question remains: *how* is the mind joined to (part of) the body? What exactly is the special relationship between the mind and the (part of the) body to which it is "joined?" On the standard interpretation of Descartes's position, he had a definite answer to this question. This is that mind and body are *causally* related. More specifically, Descartes is usually interpreted as having held that there is a two-way causal relation between mind and body: (1) the mind

causally affects the body, and (2) the body causally affects the mind. Mind-to-body causation occurs especially in voluntary action; for example, your willing or deciding to raise your arm is an act of mind that causes a physical occurrence: your arm goes up. Body-to-mind causation occurs especially in sense perception; for example, a clap of thunder is a physical occurrence that causes physical changes in your ears, nerves and brain, which in turn cause a conscious, auditory experience in your mind (which, incidentally, must on the dualistic view obviously not be confused with your brain). This theory of the mind-body relationship is called “dualistic interactionism” (where “dualistic” refers to the “twoness” or duality, of mind and matter, and “interactionism,” to the causal interaction between the two).

Dualistic interactionism has an undeniable appeal, for at least two reasons. First, it seems to harmonize very well with our experience: it certainly seems to us that our decisions and volitions frequently cause our bodies to behave in various ways—that our minds do, to an important extent, “control” our bodies. And it seems just as obvious that what happens to our bodies causes a multitude of different conscious experiences in our minds. Mind-body interaction thus seems to be a continuing and pervasive feature of our ordinary experience. Second, dualistic interactionism provides a plausible way to understand the close relationship that each of us has to his or her own body and to no other body. What makes a certain body *my* body is that *I* have a direct control over it that no one else has and that what happens to it has a direct effect on me that it has on no one else. Only I can cause that body to move by a mere volition; only I will feel pain if that body is injured. The body which I call mine, then, is the body over whose movements I have direct control and whose vicissitudes have a direct effect on me.

Despite its initial appeal, however, dualistic interactionism also faces deep difficulties. The most striking and famous one surfaces as soon as we ask the following question: *How can the mind causally affect the body, and vice-versa?* Remember, again, that according to Descartes the mind is a thinking substance that has no spatial dimensions, while body is a thoughtless, three-dimensional substance. How do these two things interact? We obviously cannot say that the one makes contact with or pushes the other, for this would require that they both have spatial dimensions and spatial surfaces. We cannot say that the mind applies a physical force to the body, since force is mass times acceleration and the mind has no mass. We cannot say that the body imparts thoughts to the

mind, since the body has no thoughts. When we actually try to conceive the interaction between an extended nonthinking thing and a nonextended thinking thing, it seems quite inconceivable—as inconceivable as driving a nail with an “immortal” hammer, or denting a spirit with rock.

In fairness, we should note that Descartes tried to give an account of mind-body interaction. He did so in the texts where he says that the mind is joined to a particular part of the brain. There he specifies the part of the brain in question, and describes the immediate consequences of its interaction with the mind. In *The Passions of the Soul*, he says:

[T]he part of the body in which the soul directly exercises its functions is not the heart at all, or the whole of the brain. It is rather the innermost part of the brain, which is a certain very small gland situated in the middle of the brain's substance and suspended above the passage through which the spirits in the brain's anterior cavities communicate with those in its posterior cavities. The slightest movements on the part of this gland may alter very greatly the course of these spirits, and conversely any change, however slight, taking place in the course of the spirits may do much to change the movements of the gland. (CSM I 340, SPW 230, AT XI 352)

The gland to which Descartes is here referring is called the pineal gland (also called the conarium). Descartes believed that this tiny gland must be the point where mind and body interact, because it is the only part of the brain that does not have a “double” (CSMK 143, 162). He theorized that it is surrounded by a refined material substance called “animal spirits,” which interact via tube-like nerves with the muscles that control various parts of our bodies. The pineal gland itself, he thought, interacts directly with the mind: a given event in the mind—say, the willing to raise one’s arm—moves the gland, causing it to drive the animal spirits through the nerves to the muscles, which then contract, thus raising the arm. Conversely, a given bodily event—say, stimulation of the retina—drives the animal spirits through the nerves to the pineal gland, whose oscillation then affects the mind in such a way that it has a certain visual experience. As Descartes put it:

[T]he small gland which is the principal seat of the soul is suspended within the cavities containing these spirits, so that it can

be moved by them in as many different ways as there are perceptible differences in the objects. But it can also be moved in various different ways by the soul, whose nature is such that it receives as many different impressions—that is, it has as many different perceptions as there occur different movements in this gland. And conversely, the mechanism of our body is so constructed that simply by this gland's being moved in any way by the soul or by any other cause, it drives the surrounding spirits towards the pores of the brain, which direct them through the nerves to the muscles; and in this way the gland makes the spirits move the limbs. (CSM I 341, SPW 231, AT XI 354–5)

Although the scientific details of this “pineal gland theory” are obsolete, the theory is nonetheless very instructive; for it provides a vivid, concrete illustration of two points. First, the dualistic interactionist who holds that the mind interacts with the brain must say that some such account as Descartes's is literally true, even if the details of Descartes's own account are wrong. In other words, the interactionist is committed to the view that certain specific brain events cause what we may call “mental events” and that certain specific mental events cause certain specific brain events. Second, such causal interaction of brain events and mental events is extremely hard to comprehend. In other words, even if we assume that the specific events involved have been correctly identified, the mystery of *how* the mental events cause physical ones, and vice-versa, remains. John Cottingham puts the problem this way:

What strikes the reader here is not so much the wealth of obsolete physiological detail (modern readers will readily be able to substitute electrochemical events in the cerebral cortex for Descartes's movements of the pineal gland and ‘animal spirits’) as the way in which that physiological detail is expected to ‘mesh’ with events in the nonphysical realm of the soul. Descartes has managed to supply a host of mechanisms whereby movements, once initiated in the pineal gland, can be transferred to other parts of the brain and body; but he does not seem to have tackled the central issue of how an incorporeal soul can initiate such movements in the first place. And the same problem will apply when the causal flow is in the other direction. Descartes devotes a lot of attention to the physiological mechanisms whereby bodily

stimuli of various kinds cause changes in the nervous system and brain which ‘dispose’ the soul to feel emotions like anger or fear. But he does not explain how mere brain events, however complex their physiological genesis, could have the power to arouse or excite events in the mental realm.²⁶

Another recent writer, Fred Feldman, gives an illustration that makes the mystery seem even worse when it is put in more up-to-date terms than Descartes’s pineal gland theory:

[S]uppose someone dips his toe into a swimming pool to test the water. The cold water quickly cools the skin on his toe, and changes the temperature of the nerve endings that are “scattered” there. Then some sort of electrical charge flows up the nerve, jumping across various gaps between one nerve and the next. Perhaps there are stages in which the electrical event causes some chemical change, which in turn causes a suitable electrical event in the next nerve. This purely physical chain of events eventually reaches some part of the brain. Here is where the trouble begins. How does it make the last step, the one that gets it from the physical apparatus of the nervous system, and into the mind? . . . How does the body finally influence the mind? How does all the electrical and chemical activity in the nervous system finally bring about that distinctive feeling of cold that reveals that the water is too chilly for swimming? Many philosophers would say that this alleged causal connection is simply inconceivable.

The causal connection in the other direction is no easier to understand. Suppose you have been thinking about a certain friend. You decide to call her on the telephone. Precisely how does this decision, apparently a mental event, give rise to the first physical event in the causal chain that ultimately leads to the movements of your fingers? Your mind cannot rub against the nearest nerves in your brain, nor can it give off heat, or light, or chemicals, or an electric charge. Only a physical object could do such things. Thus, causal interaction is equally perplexing, whether it is mind-body interaction, or body-mind interaction.

²⁶ Cottingham, “Descartes,” p. 223.

And Cartesian dualism cannot be true unless such interaction happens all the time. . . . [T]his . . . constitutes the most important objection to Cartesian dualism. . . . [H]istorically, . . . no other argument crops up more frequently in the anti-dualistic literature. It is a formidable problem.²⁷

Indeed, mind-body interaction has, from Descartes's own day to ours, often been seen as the weakest point of his philosophy. Princess Elizabeth of Bohemia, with whom Descartes had an important philosophical correspondence, wrote to him that she found it easier to conceive that the mind itself is extended, than to conceive how it interacts with extension. And Gassendi, in his *Objections to the Meditations*, powerfully attacked the idea of interaction (see, e.g., CSM II, 235–239, AT VII 339–345). Still other contemporaries of Descartes proposed various ways of amending his system so as to avoid interaction altogether. For example, the Flemish thinker Arnold Geulincx (1624–1669) and the French philosopher Nicolas Malebranche (1638–1715) proposed a theory called *Occasionalism*. According to this theory, there is no interaction between mind and body. Instead, on each separate occasion where mind-body interaction *seems* to take place, it is really God who brings about the effect. So, for instance, when you decide to raise your arm, God himself “steps in” and raises it for you; when you accidentally touch a hot stove, God produces pain in you. The great German rationalist Leibniz proposed a somewhat different theory, called *Preestablished Harmony*. Leibniz agreed with the Occasionalists that there is no interaction between mental and physical events. However, he thought it unbefitting to God’s perfection that he should need to constantly intervene so as to coordinate these events. Accordingly, he theorized that God, when he created the world, also instituted a universal correspondence or harmony between mental and physical events, somewhat like a skillful clockmaker who builds two clocks in such a way that they will always keep exactly the same time. Today, philosophers who work on the mind-body problem regard both Occasionalism and the Preestablished Harmony as theories that are too fantastic to warrant serious consideration. We mention them here because they vividly show how radical Descartes’s problem of interaction seemed to be to his own contemporaries. The problem seemed so intractable that even theories that made God the intermediary between the mind and body of a person were seen by leading thinkers of the day as genuine improvements over interactionism.

²⁷ Fred Feldman, *A Cartesian Introduction to Philosophy*, pp. 202–3.

Indeed, some scholars have suggested that Descartes himself was ultimately drawn to occasionalism. Any attempt to defend an occasionalist reading of Descartes, however, would have to contend with passages like the following, where he seems unequivocally to affirm mind-body interaction:

That the mind, which is incorporeal, can set the body in motion is something which is shown to us not by any reasoning or comparison with other matters, but by the surest and plainest everyday experience. It is one of those self-evident things which we only make obscure when we try to explain it in terms of other things. (Letter to Arnauld, July 29, 1648; CSMK 358)

It is not necessary for the mind to be a body, although it has the power of moving the body. (*Fifth Replies*, CSM II 266, AT II 389)

[Burman] But how can this be, and how can the soul be affected by the body and vice-versa, when their natures are completely different? [Descartes] This is very difficult to explain; but here our experience is sufficient, since it is so clear on this point that it just cannot be gainsaid. (*Conversation with Burman*, CSMK 346, AT V 163)

I would say: ‘. . . the brain alone can directly act on the mind.’ (Letter to Regius, May 1641, CSMK 183, AT III 373)

Is the problem of interaction a fatal difficulty for dualistic interactionism? Although the difficulty is clearly a major one, there is reason to think that it is not insuperable. For some recent dualists (e.g., Broad and Ducasse) have suggested ingenious ways of dealing with it. Broad argued that the problem arises from the assumption that cause and effect must be similar, and that this assumption is false even in the purely physical realm: for example, a magnetic force can attract iron filings, but there is no resemblance between the magnetic field and the motions of little pieces of iron.²⁸ To this, it may be countered that a magnetic force and the motions of iron filings are at least both physical phenomena that can be investigated by physics, whereas this is not the case for mind and body. But Ducasse’s reply to the difficulty is especially incisive. He writes:

²⁸ C. D. Broad, *Mind and its Place in Nature*, p. 97. Broad’s own example is that of a draft causing a cold.

[T]he objection that we cannot understand how a psychical [i.e. mental] event could cause a physical one (or vice versa) has no basis other than blindness to the fact that the “how” of causation is capable at all of being either mysterious or understood only in cases of *remote causation*, never in cases of proximate causation. For the question as to the “how” of causation of a given event by a given other event never has any other sense than *through what intermediary causal steps* does one cause the other.²⁹

We can illustrate Ducasse's point with an example. Suppose someone is riding a bicycle, and that it is asked, “How does pushing on the pedals cause the bicycle to move?” The answer is that pushing on the pedals causes them to move, which causes the front sprocket to turn, which causes the chain to rotate, which causes the rear wheel to turn, which causes the bicycle to roll forward. In giving this answer, we have specified several causal steps that are the intermediaries between the pushing of the pedals and the movement of the bicycle. In specifying these intermediary causal steps, we have explained *how* pushing the pedals causes the bicycle to move, or *how* the cause produces the effect. But now suppose somebody were to ask: “But how does pushing on the pedals cause them to move?” The answer is that it just does. We cannot explain how it does, because there are no intermediary causal steps between the cause and the effect that we could cite in order to answer this “how” question. We have reached a case of *proximate causation*, one where there are no intervening steps between the cause and the effect; and such cases cannot be explained. Instead, they have to be simply accepted as brute facts: “That's just how the world works.”

How, exactly, does Ducasse's point bear on the problem of interaction? The answer is that dualistic interactionists have always taken mind-body interactions to be cases of proximate causation. Therefore, the question of how a given mental event causes a brain event—or how a given brain event causes a mental event—is just as illegitimate as the question of how pushing on the pedals causes them to move. To make this more concrete, suppose (as C. D. Broad suggested in his book *Mind and Its Place in Nature*) that the immediate effect of a mental event on the brain is to lower the electrical resistance of certain synapses and to raise that of

²⁹ C. J. Ducasse, “In Defense of Dualism,” p. 88.

others, so that the directions of the paths taken by nervous impulses through those synapses are affected. And suppose it is asked, "How does the desire to raise one's arm cause it to go up?" The answer, in outline, is that this mental event alters the distribution of resistances in certain synapses, which affects the direction of the paths taken by certain nervous impulses, which causes certain muscles to contract, which causes one's arm to go up. But now, if someone asks how desiring to raise one's arm causes the electrical resistance in some synapses to go down and in others to go up, then the only answer that can be given is that this is a case of proximate causation. As such, it cannot be explained. It must be simply accepted as a brute fact.

At this point, you may want to raise the following objection. Surely, when a bicycle rider causes the pedals to move by pushing on them, there is an explanation of their motion, in terms of laws of nature involving force, momentum, transfer of energy, or the like. The interactionist cannot sensibly deny this. But then, how can he say that such cases of proximate causation are just inexplicable brute facts?

A possible reply is that the interactionist need not at all deny that laws of nature, involving various scientific concepts, govern cases of proximate causation. But this does not mean that the proximate causal connections are any the less brute facts. For what are these laws of nature? According to one important and plausible view, they are basically *regularities* between events, to the effect that whenever an event of a given kind occurs, then an event of another kind occurs. For instance, whenever a given force is applied to a body (under certain specifiable conditions), it moves. The key point to grasp here is that the law merely states that the case in question is an instance of what always happens in similar cases. It does not explain *how* applying a force to a body (e.g., pushing on a bicycle pedal) makes it move, except in the sense of citing a *regularity* in nature, to the effect that whenever such a force is applied, the body moves. This regularity itself, and particular cases of the regularity, remain as much brute facts as ever. The most we can do by way of explaining the regularity is to derive it from a still more general regularity. But then the latter is itself a brute fact: it is just a very general principle describing the way nature happens to behave. All that the interactionist is claiming is that, among the regularities in nature, there are some to the effect that whenever a certain kind of mental event occurs, a certain kind of brain event occurs; and some to the effect that

whenever a certain kind of brain event occurs, a certain kind of mental event occurs. Such regularities are no more nor less mysterious than those linking purely physical events.

It must be acknowledged that the “regularity” conception of a law of nature just sketched, which derives from David Hume’s famous and influential analysis of causality, is not uncontroversial. It is, however, one of the leading philosophical theories about laws of nature. To the extent that it is plausible, it supports Ducasse’s solution to the problem of interaction (though we should note that Ducasse himself did not accept Hume’s theory and so would probably not have approved of our invoking it to support his solution to the problem).

6. An Assessment of Cartesian Dualism

In this chapter’s first section, we presented Descartes’s proof of dualism and Arnauld’s objection to it and argued that Descartes gives a plausible reply to that objection. In the last section, we presented the problem of interaction and argued that this problem does not constitute a decisive objection to dualistic interactionism. Does this mean that we should accept Cartesian Dualism (i.e., Descartes’s entire theory of mind and matter as two different but causally interacting substances)? No: that would be a hasty conclusion for at least three reasons. First, although we have discussed the most famous objection to dualistic interactionism, there are other objections to it that we have not considered. Second, there are other dualistic theories that we also have not considered, such as *epiphenomenalism*, a theory maintaining that physical events cause mental events but that mental events never cause physical events. Third, many (perhaps most) contemporary philosophers reject all forms of dualism in favor of dualism’s main rival, which is *materialism*. Materialism is the view that whatever exists is material. According to materialism, there are no such things as mental substances, mental events, or mental states. If “minds” exist at all, then they can only be brains, or brains together with nervous systems. If thoughts, desires, feelings and sensations exist (something which is actually denied by the most radical materialist philosophers, called “eliminative materialists”), then they are merely neurological events, states or processes occurring within the brain and central nervous system. Today, there are many

different versions of materialism, with different names such as “eliminative materialism,” “logical behaviorism,” “reductive materialism,” and some versions of “functionalism.”

We shall not examine materialist theories of mind, nor shall we examine in detail other objections to dualistic interactionism or other dualistic theories like epiphenomenalism. It is not that these topics are unimportant or uninteresting. On the contrary, the area of philosophy to which they belong, called Philosophy of Mind, is a very active and exciting field. Our reason for not going more deeply into the philosophy of mind is simply that doing so would take us too far from our study of Descartes and would indeed require another book.

Nevertheless, we shall conclude this book by offering an overall assessment of Cartesian Dualism. In the course of our discussion, we shall touch on a few of the issues raised in contemporary philosophy of mind. But let us start with a note of caution: no final judgment as to the truth of Cartesian Dualism can reasonably be made apart from the detailed examination of other objections and rival theories that we have decided to forego. So, our assessment will not take the form of pronouncing upon the truth or falsity of the theory. Rather, we shall defend the more modest thesis that even if Descartes’s proof of dualism is sound, this does not have the weighty implications that are often associated with Cartesian Dualism—implications prized by dualists and shunned by materialists.

To defend this thesis, we shall use a four-point strategy. First, we shall present a pruned, or streamlined, version of Descartes’s proof of dualism that preserves the basic intuitions or insights on which Descartes’s own argument turns. Second, we shall argue that even if this streamlined argument is sound, it does not have the significant implications commonly associated with dualism. Third, we shall argue that appearances to the contrary notwithstanding, these implications do not follow from Descartes’s own argument, either—that his argument does not have weightier implications than the streamlined one. Finally, we shall show that Descartes was aware of this—that he was not fooled by his argument’s impressive look.

Descartes’s proof of dualism turns on three basic ideas: the conceivability of the separate existence of the mental and the physical, the principle that what is conceivable is logically possible, and the principle that the logical possibility of separate existence implies distinctness or nonidentity. Thus, our streamlined version of his argument starts from the premise

- (1) If X is any conscious state (i.e., any thought, desire, mental image, sensory experience, or the like) and Y is any physical state or process (e.g., a brain event), then it is conceivable that X exists and Y does not exist.

This premise resembles, and rests on the same considerations as, Descartes's richer premise (stated as line 5 of his proof of dualism in section 2 of this chapter) that he can clearly and distinctly conceive himself, as a thinking and nonextended thing, existing apart from his body, as an extended and nonthinking thing. For brevity's sake, we have used the term "conceivable," instead of "clearly and distinctly conceivable," in stating premise 1 of the streamlined argument. But of course, the qualification, "clearly and distinctly," should be understood in (1) and throughout the rest of the streamlined argument. The second premise of that argument is

- (2) If it is conceivable that X exists and Y does not, then it is logically possible that X exists and Y does not.

This premise is very close to Descartes's premise (stated as line 2 of his proof of dualism) that if he can clearly and distinctly conceive X existing apart from Y, then X really can exist without Y, at least by God's power. The third premise is

- (3) If it is logically possible that X exists and Y does not, then X is not identical with Y.

This is close to Descartes's premise (stated as line 3 of his proof of dualism) that if X can really exist without Y, no matter by what power, then X and Y are really two different things. The basic point, as before, is that nothing could possibly exist without or apart from itself. Of course, neither (2) nor (3) of the streamlined argument makes any reference to God's power. For Descartes's philosophical theology, as previously suggested, plays no essential role in his proof of dualism (unlike his proof of the physical world). Rather, the key ideas in the proof of dualism are the three listed at the start of this paragraph, each of which is expressed in just one of the streamlined argument's three premises. The conclusion that follows from those premises is

- (4) If X is any conscious state and Y is any physical state or process, then X is not identical with Y.

This conclusion is weaker than Descartes's own conclusion that he, as a thinking thing, is not identical with his body. For one thing, (4) asserts only a dualism of *states or processes*, as opposed to a dualism of *substances*. For another, (4) contains no reference to "I," or to the *self*. Below, when we come to our point that Descartes's own argument does not have significantly weightier implications than does our streamlined argument, we shall argue that these differences do not matter very much. For the moment, however, we need only emphasize that (4) does, of course, assert a *dualism* of the mental and the physical; for it implies (assuming that any conscious states and physical states exist at all) what all materialists deny, namely, that in addition to material things and processes, there are purely mental existents. Since (4) asserts a dualism of states or properties rather than of substances, it is the minimum that a dualist must hold; it is a minimal form of dualism—but (4) is still dualism, since it means that mental states are irreducibly different from anything physical. It also directly contradicts one of the most widely-held recent versions of materialism, the "identity theory," also called reductive materialism. This theory doesn't deny that there are mental states, but it just baldly identifies these with brain states or processes. Thus the streamlined argument brings Cartesian considerations to bear directly against a major form of contemporary materialism.

Our streamlined version of Descartes's proof of the Real Distinction can doubtless be attacked. For example, some philosophers today would question (2), for (2) rests on the general principle that whatever is (clearly and distinctly) conceivable is logically possible. But these philosophers think that there are counter-examples to this principle. One favorite example is that of Goldbach's Conjecture. Christian Goldbach (1690–1764) was a Russian mathematician who conjectured that every even number greater than 2 is the sum of two prime numbers. So far, every such number which has been "tested" has been found to be the sum of two prime numbers; but, of course, this does not mean that some very large even number that no one has yet considered isn't the sum of two primes. Furthermore, no mathematician has been able either to prove or to disprove Goldbach's Conjecture. On the strength of this example, some philosophers argue that the general principle behind (2) is false. The argument is that since Goldbach's Conjecture is a statement of mathematics, it is logically necessary if true at all, and logically impossible if false. But we can conceive of Goldbach's Conjecture being either true or false (i.e., we can conceive of both

alternatives), say, by conceiving that some supercomputer tells us that it is true or that it is false. Therefore, we can conceive of something that is not even logically possible: the truth of Goldbach's conjecture (in case the conjecture is actually false), or the falsity of the conjecture (in case it is actually true). So, the principle that whatever is conceivable is logically possible is false. It seems, however, that this objection commits a rather simple mistake. It is one thing to conceive of a computer telling us that Goldbach's conjecture is true (or false) and quite another to conceive that the conjecture itself is true or false. I can conceive of a computer telling me, $1 + 1 = 3$; it does not follow that I can conceive $1 + 1 = 3$. As Descartes would put it, such a "conception" would hardly be clear and distinct. It seems, therefore, that the principle behind (2)—a principle that many philosophers rely upon—cannot be so easily dismissed.

We need not insist, however, that the streamlined argument can withstand this or other objections that might be raised against it; for although the streamlined argument seems quite powerful, our point is not that it is sound. Rather, our point (the second one in the four-point strategy outlined above) is that even if the streamlined argument is sound, the implications are not very significant. The basic reason for this is that even if the logical possibility that conscious states might exist apart from physical states shows that conscious states are not identical with physical states, *it does not show that it is causally possible for conscious states to exist apart from physical states*. Thus, for all that the argument can show, consciousness may exist only as an effect of certain sorts of brain processes; that is, it may never exist apart from the physical processes that cause it. This may be so even if (as interactionism maintains but epiphenomenalism denies) mental events causally affect brain processes or events; for it may be that those mental events are in turn caused by prior brain events and never occur unless caused by these brain events. A contemporary philosopher, the late James Cornman, drawing on Broad's suggestions about mind-body interaction, suggested one way in which this might happen. Cornman suggests that the causal relationships between mental and brain events could be as shown in Figure 6.1.

If this is the way mental and brain events are interrelated, then there are two notable consequences. First, as Cornman emphasizes, a scientist could give an adequate neurological explanation of human thought and behavior without ever mentioning the mental events; for although M causally affects S, M is in turn caused by B. So a neurophysiologist who

Figure 6.1

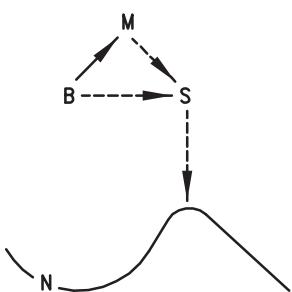
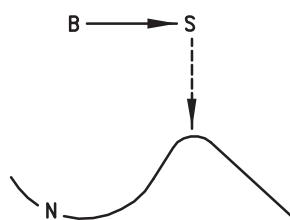


Figure 6.2



B = brain event; M = mental event; N = nerve impulse; S = change of electrical resistance in synapses; $x \rightarrow y$ = x causes y; $x \dashrightarrow y$ = x causally affects y; = path of nerve impulse.

Source: Slightly modified from James W. Corman, Keith Lehrer, and George Pappas, *Philosophical Problems and Arguments: An Introduction*, 3d ed. (Indianapolis and Cambridge: Hackett Publishing Company, 1987), p. 172. Chapter 4 of this book contains an excellent discussion of the various objections and alternatives to dualistic interactionism.

knew that B was occurring could predict and explain S and N without ever mentioning M—even, indeed, if the neurophysiologist were a materialist who denied M's existence and believed that matters stand as in Figure 6.2. Second, since on this view M would not occur unless B (or some other brain event) produced it, consciousness is totally dependent on a functioning brain for its existence.

Once these points are recognized, some of the objections that recent materialists raise against dualism seem strangely empty and rhetorical. For example, one leading contemporary materialist's chief objection to dualism is that it postulates "an irreducibly psychical something," a "ghost stuff . . . or ripples in an underlying ghost stuff," different from anything recognized by the natural sciences, thereby violating the principle of theoretical simplicity.³⁰ (The principle of theoretical simplicity, also known as Ockham's Razor, is a very general methodological principle saying that a theory should not "multiply entities" unnecessarily, that is, should not postulate more different kinds of entities than are needed to explain the facts that the theory is designed to explain. A very simple illustration would be that since tornadoes can be explained in terms of certain atmospheric conditions, a meteorological theory should not postulate evil spirits to explain their occurrence. A historical example of the principle's

³⁰ J. J. C. Smart, "Sensations and Brain Processes," p. 53, p. 64. in C. V. Borst, ed., *The Mind-Brain Identity Theory* (New York: St. Martin's Press, 1970), pp. 52–66.

application would be the abandonment of the phlogiston theory after the discovery of the role of oxygen in combustion.) But if dualism concedes that mental existents may be completely causally dependent on physical ones, then it is hard to see much force in this complaint. To be sure, if the argument for dualism is sound, then conscious states cannot be reduced to brain states: in that sense they are indeed irreducibly mental or psychical. But phrases like "irreducibly psychical something," "ghost stuff," and the like also suggest a kind of independence of mental existents from physical ones which is wholly belied by such a dualism. And while the notion of theoretical simplicity is a complex one, its applicability here is questionable; for it seems reasonable to hold that its sphere of application is limited to entities that are held to play some essential *explanatory* role (especially in scientific explanation), though there is no independent evidence for their existence. But then the principle of simplicity cannot be legitimately invoked to rule out a mental event like M, since M plays no essential explanatory role and there is independent evidence for its existence (i.e., the streamlined argument together with the evidence of introspection).

The other side of the coin, of course, is that if the dualist concedes that the mental may be completely causally dependent on the physical, then it is hard to see the great significance of dualism; for such a dualist agrees with contemporary materialists who think that a complete neurophysiology could causally explain human thought and behavior and that there is good reason to believe that consciousness ends when brain activity stops. Such a dualist insists only on the logical possibility of consciousness existing apart from brain activity and its consequent nonidentity with brain activity. This is to abandon some of the ideas commonly associated with dualism, for example, that the mental component of human beings stands in the way of any adequate scientific explanation of human thought and behavior and offers some positive evidence for immortality. Thus, the objections of materialists are blunted only at the cost of foregoing some of the most cherished ideas commonly associated with dualism. The disagreement between dualists and materialists seems to become at least partly a verbal one, over labels like "spiritual" and "material," and associated ideological beliefs that do not logically follow even if the case for dualism is sound.

Appearances to the contrary notwithstanding, much the same assessment holds for Descartes's own proof of dualism. This may at first not be apparent, because Descartes's proof of the Real Distinction between mind and body may seem to yield a much stronger form of dualism than does the streamlined argument. But this is a misleading impression.

First of all, remember that Descartes's argument does not try to show that the mind does, at any time, exist without the body. Rather, as we saw in Section 2, it only tries to show that the mind *could* exist without the body. As we saw, Descartes holds that the mere logical possibility that mind can exist apart from body is enough to show that they are different things, or that there is a real distinction between them.

But even in light of a proper grasp of the Real Distinction, it may still seem that Descartes's argument has weightier implications than our streamlined version of it. There are two reasons for this. The first is that Descartes's own argument is intended to establish a dualism of substances, rather than merely a dualism of states. The second is that Descartes calls one of these substances "I," or *self*, and says that it constitutes his "essence," thereby giving it a special priority over the other.

Neither of these points, however, is as significant as it may seem. Let us consider the second point first. Suppose we grant that Descartes's argument shows that it is logically possible that I exist without my body and that it is logically possible that I exist as only a mind, so that my mind's existence is logically sufficient for my existence. Suppose we also agree that it is not logically possible that I exist without my mind, that is, that my mind's existence is logically necessary for my existence. Finally, suppose we agree that "I" should denote only that which is both logically necessary and sufficient for me to exist. It then follows that "I" denotes my mind, not my body. It also follows, by the principle that a thing's essence comprises whatever is necessary and sufficient for its existence, that my essence is my mind, not my body. To put it more formally, suppose that we grant the entire argument:

1. My mind's existence is logically sufficient for my existence.
2. My mind's existence is logically necessary for my existence.
3. "I" denotes only what is both logically necessary and logically sufficient for my existence.
4. "I" denotes my mind and not my body (from 1, 2, and 3).
5. A thing's essence contains only what is both logically necessary and logically sufficient for its existence.
6. My essence is my mind, not my body (from 1, 2, and 5).

Now even if we accept this reasoning completely, it only establishes a point of logical, linguistic, or conceptual priority. It has no tendency whatsoever to show that what "I" denotes, or that what is comprised in my essence, is *causally* independent of my body.

Consider next the fact that Descartes's own argument purports to establish a dualism of substances, rather than merely a dualism of states. The term "substance" is associated, especially within philosophy, with a certain permanence, durability, even indestructibility. This can make it look as if Descartes's argument would, if it is sound, establish a dualism that supports traditional beliefs about the causal independence of the soul from the body and its prospects for immortality. Now Empiricist criticisms of the concept of substance have cast a shadow over Descartes's reliance on this concept. These criticisms also go a long way toward explaining why, today, dualism is usually formulated in terms of mental versus physical states or events, rather than mental versus physical substance. But the important point for us is that even if dualism is formulated in terms of mental and physical substance, our basic assessment of Descartes's proof of dualism still holds good; for the argument that Descartes gives for the substance theory, namely the argument from change, certainly does not show that a mental substance would be causally independent of the body, or that it would be indestructible, or that it must outlast the body, or anything of the sort. It follows that nothing in the concept of substance, at least insofar as its employment is supposed to be justified by this argument, entails that a mental substance and its states must be causally independent of a material substance and its states. Further, there do not seem to be any other arguments concerning substance that would support such a conclusion. Certainly, the proof of the Real Distinction contributes nothing itself that might strengthen the concept of substance; it only uses the concept of substance that was already operative in *Meditation II*. So far as Descartes's argument for dualism goes, then, mind might be totally dependent on matter.

Indeed, this would be the case if mental states or events were completely causally dependent on physical ones. To see this, recall once again that just as there cannot be a property without a substance to which it belongs, so there cannot be a substance without any properties. It follows that a mental substance cannot exist unless it has properties. Now the properties of a mental substance must be its various "thoughts"—its mental states. Thus, a mental substance cannot exist unless it has mental states. So, if those mental states are totally causally dependent on physical states or events (e.g., on brain states or brain events), then it follows that mental substance itself depends for its existence on the occurrence of these physical states or events. Furthermore, since those physical states are themselves properties that cannot exist apart from a material substance, it also follows that mental substance depends for its existence

on material substance. Putting this a bit more formally, we can give the following argument. The substance theory, combined with the view that a mental substance must have the appropriate sort properties (namely, mental properties), implies that

- (1) If there are no instantiated mental properties (= mental states), then there are no mental substances.

But the causal dependence of mental states on physical states, which is not ruled out by Descartes's proof of the Real Distinction, means that

- (2) If there are no instantiated physical properties, then there are no instantiated mental properties.

From (1) and (2), it follows that

- (3) If there are no instantiated physical properties, then there are no mental substances.

Now the substance theory, combined with the view that physical properties can belong only to the appropriate kind of substance (namely, material substance) implies that

- (4) If there are no material substances, then there are no instantiated physical properties.

Finally, there follows from (3) and (4) that

- (5) If there are no material substances, then there are no mental substances.

The point can also be shown schematically, as in Figure 6.3 below.

Nothing in Descartes's case for dualism rules out such dependence of *res cogitans* on *res extensa*. At best Descartes's arguments give a certain *epistemological* priority to *res cogitans*—show that its existence can be known before that of *res extensa*. But this does not mean that *res cogitans* has any *metaphysical* priority—that it can actually exist independently of *res extensa*.

Did Descartes himself recognize the limitations of his case for dualism? It might seem that he did not; for in principle 51 of the *Principles*

abbreviations:

X → Y = X depends on Y

(E) _____ = existence of _____

(I) _____ = instantiation/occurrence of _____

(E) mental substance → (I) mental properties

(I) mental properties → (I) physical properties

(I) physical properties → (E) physical substance

∴ (E) mental substance → (E) physical substance

Figure 6.3

of *Philosophy* he writes, “By *substance* we can understand nothing other than a thing which exists in such a way as to depend on no other thing for its existence” (CSM I 210, SPW 177, AT VIII A 24). At first sight, this definition seems to contradict flatly our claims about the possible causal dependence of mental substance on material substance. However, in the next sentence Descartes says:

And there is only one substance which can be understood to depend on no other thing whatsoever, namely God. In the case of all other substances, we perceive that they can exist only with the help of God’s concurrence. Hence the term ‘substance’ does not apply *univocally* [= with one and the same meaning] . . . to God and to other things; that is, there is no distinctly intelligible meaning of the term which is common to God and his creatures. <In the case of created things, some are of such a nature that they cannot exist without other things, while some need only the ordinary concurrence of God in order to exist. We make this distinction by calling the latter ‘substances’ and the former ‘qualities’ or ‘attributes’ of those substances.> ³¹

³¹ The last sentence is in angle brackets because Descartes added it in French to the original Latin text of the *Principles*.

Here, Descartes admits that the definition of substance he has just given applies only to God. (Later, Spinoza was to argue that Descartes should have stuck to this definition and so embraced Spinoza's own view that God is the only substance that exists or can be conceived to exist.) Created substances cannot be defined as things that need nothing else in order to exist, because they require God's constant concurrence in order to exist. They qualify as "substances" only in the weaker sense that they need nothing else *except God's concurrence* in order to exist. As Descartes says in principle 52, "as for corporeal substance and mind (or created thinking substance), these can be understood to fall under this common concept: things that need only the concurrence of God in order to exist" (CSM I 210, SPW 177, AT VIIIA 25). This need for God's concurrence stems from Descartes's doctrine, introduced in *Meditation III*, that created substances, both mental and physical, need to be "continuously created" by God in order to remain in existence (CSM II 33, SPW 96, AT VII 48–9). Although we need not examine this doctrine (which we briefly noted in chapter 3) for its own sake, we should take it into account in considering whether Descartes thought that his case for dualism rules out the causal dependence of one created substance on another. In alluding to the doctrine in the *Principles of Philosophy*, Descartes does not explicitly differentiate between saying that (a) created substances exist *only if* God sustains them, and (b) created substances exist *if only God sustains them* (in other words, if and only if God sustains them). Thus, Principle 51 says that "other [i.e., created] substances . . . can exist only with the help of God's concurrence," while Principle 52 says that created substances are "things that need only the concurrence of God in order to exist" (statements (a) and (b), respectively). The former is consistent with one created substance's causally depending on another (as well as on God), whereas the latter is not.

Now it seems that Descartes must have held (b) as well as (a). For surely, an omnipotent God could, if he wished, sustain a substance in existence without the help of any other thing. However, it does not follow that Descartes has shown—or that he believes he has shown—that the mind *is* a substance in the sense stipulated by (b). For all that the proof of the real distinction has shown, it *might* be that a human mind, like a "quality" or "attribute" is one of those "created things, [that has] . . . such a nature that they cannot exist without other things." The world might work in such a way—God might have designed it in such a way—that human minds, even though an omnipotent being could of course

(miraculously) sustain them in existence without the help of anything physical, naturally depend on physical things and, therefore, do not count as substances even in the weaker sense that applies to created things. That Descartes has not shown the contrary of this should be obvious from what has already been said about his rationale for using the concept of substance. More interestingly, there is excellent evidence that Descartes realized he had not shown it. In the *Synopsis of the Meditations*, he issues a warning to “people [who] may perhaps expect arguments for the immortality of the soul” in the *Meditations* (CSM II 9, SPW 73, AT VII 12). He explains that since his aim in the *Meditations* was to assert only what he could rigorously prove, he was obliged to follow strictly his method of never asserting a proposition before setting out all the premises on which it depended. He then declares that

the first and most important prerequisite for knowledge of the immortality of the soul is to form a concept of the soul which is as clear as possible and is also quite distinct from every concept of body; and that is just what has been done in this section [in *Meditation II*]. (CSM II 9, SPW 73, AT VII 13)

Next, he outlines the further steps that will finally lead up to his assertion of the Real Distinction in *Meditation VI*. Then he adds this crucial comment (most of which we have already quoted in discussing other points):

But I have not pursued this topic [i.e., immortality] further in this book, first because these arguments are enough to show that the decay of the body does not imply the destruction of the mind, and are hence enough to give mortals the hope of an after-life, and secondly because the premises which lead to the conclusion that the soul is immortal depend on an account of the whole of physics. This is required for two reasons. First, we need to know that absolutely all substances, or things which must be created by God in order to exist, are by their nature incorruptible and cannot ever cease to exist unless they are reduced to nothingness by God’s denying his concurrence to them. Secondly, we need to recognize that body, taken in the general sense, is a substance, so that it too never perishes. But the human body, in so far as it differs from other bodies, is simply made up of a certain

configuration of limbs and other accidents of this sort; whereas the human mind is not made up of any accidents in this way, but is a pure substance. For even if all the accidents of the mind change, so that it has different objects of the understanding and different desires and sensations, it does not on that account become a different mind; whereas the human body loses its identity merely as a result of a change in the shape of some of its parts. And it follows from this that while the body can very easily perish, the mind is immortal by its very nature. (CSM II 10, SPW 74, AT VII 13–14)

There are two important points to notice about what Descartes says here. First, he represents the thesis that the “absolutely all substances, or things which must be created by God in order to exist” are naturally incorruptible (i.e., that they need only God’s concurrence in order to exist) as *something to be established*. Notice that in calling all the “things which must be created by God in order to exist” *substances* but granting that their incorruptibility needs to be *proved*, he is using the term “substances” purely denotatively to refer to mind and body; he is not assuming that these substances even need to meet the definition of substance in his weaker sense. Thus, he is prepared to admit that for all his metaphysical arguments in the *Meditations* have shown, neither the soul nor body is incorruptible, or is a substance even in the weaker sense. Although a careless reader might think that Descartes retracts this admission at the very end of the passage, where he seems to affirm the immortality of the soul, in fact he makes no such retraction; for part of the “this” from which the soul’s immortality “follows” is the thesis, yet to be established, that “all substances . . . are by their nature incorruptible”—that is, that all the things that he has been calling “substances” really are substances in the weaker sense. In other words, Descartes here affirms the soul’s immortality only conditionally. His point is that if the thesis that the things we call substances are incorruptible were established, then, given that the soul is one of those things, *its* immortality would be established, too. The second point is even more significant. Descartes says that proving the immortality of the soul depends upon the completion of his whole *physics*. Now although it may be hard for us to see how a *physics* could show that a purely mental thing is incorruptible, the vital point is the implication that any knowledge we could have of the soul’s immortality must rest on a completed natural science, which

might reveal as-yet-unknown causal connections and dependencies between mind and body. Descartes realized that the logical possibility of the mind's existing without the body, and its consequent nonidentity with the body (which is all he had argued for in the *Meditations*) is consistent with the mind's causal dependence on the body. He knew—and openly acknowledged—how far his case for dualism was from establishing the religious beliefs that he himself undoubtedly held and cherished.

BIBLIOGRAPHY

The secondary literature on Descartes is immense. The most complete bibliographies are:

- Chappell, Vere and Doney, Willis eds. *Twenty-Five Years of Descartes Scholarship, 1960–1984: A Bibliography*. New York: Garland, 1987.
Sebba, Gregor. *Bibliographica Cartesiana: a Critical Guide to the Descartes Literature 1800–1960*. The Hague: Martinus Nijhoff, 1964.

The former has an informative introduction and cross-indexes works by topic; the latter contains abstracts of many of the works. Some shorter but useful bibliographies can be found in:

- Caton, Hiram. P. *The Origin of Subjectivity: An Essay on Descartes*. New Haven: Yale University Press, 1973.
Cottingham, John, ed., *The Cambridge Companion to Descartes*. Cambridge: Cambridge University Press, 1992.
Doney, Willis. “Some Recent Work on Descartes: a Bibliography.” In Michael Hooker, ed., *Descartes: Critical and Interpretive Essays*. Baltimore: Johns Hopkins University Press, 1978.
Gaukroger, Stephen. *The Blackwell Guide to Descartes’ Meditations*. Malden: Blackwell, 2006.
Grene, Marjorie. *Descartes*. Minneapolis: University of Minnesota Press, 1985.
Naaman-Zauderer. *Descartes’ Deontological Turn*. Cambridge: Cambridge University Press, 2010.

The list of works given below includes only the sources cited or mentioned in this book. Among the items listed, the books by John Cottingham, E. M. Curley, and Margaret Wilson contain useful bibliographies. So do Willis Doney’s first edited collection and Alan Gewirth’s “The Cartesian Circle Reconsidered.”

- Aquinas, Thomas. “Summa Theologica.” In Anton C. Pegis, ed., *The Basic Writings of Saint Thomas Aquinas*. New York: Random House, 1945.
Aristotle. “De Anima (On the Soul).” In Richard M. McKeon, ed., *The Basic Works of Aristotle*. New York: Random House, 1941.

- Berkeley, George. *A Treatise Concerning the Principles of Human Knowledge*. Oxford: Oxford University Press, 1998.
- Borst, Clive V., ed., *The Mind-Brain Identity Theory*. New York: St. Martin's Press, 1970.
- Broad, Charlie D. *Mind and Its Place in Nature*. London: Routledge & Kegan Paul, 1962.
- Burnyeat, Myles, ed., *The Skeptical Tradition*. Berkeley: University of California Press, 1983.
- Butler R.J., ed. *Cartesian Studies*. Oxford: Basil Blackwell, 1972.
- Cartwright, Richard. "Negative Existentials." *Journal of Philosophy* 57 (20/21) (1960): 629–39.
- Chappell, Vere. "Descartes's Compatibilism." In John Cottingham, ed., *Reason, Will and Sensation*. Oxford: Oxford University Press, 1994.
- Copleston, Frederick. *A History of Philosophy*, Book 2, vol. 4. New York: Doubleday, 1985.
- Cornman, James W., Keith Lehrer, and George Pappas. *Philosophical Problems and Arguments*, 3d ed. Indianapolis and Cambridge: Hackett Publishing Company, 1987.
- Cottingham, John. *A Descartes Dictionary*. Oxford: Blackwell, 1993.
- Cottingham, John. *Descartes*. Oxford: Basil Blackwell, 1986.
- Cottingham, John. "Descartes: Metaphysics and the Philosophy of Mind." In G.H.R. Parkinson, ed., *Routledge History of Philosophy*, Vol. IV. London: Routledge, 1993.
- Cottingham, John, ed., *Reason, Will and Sensation*. Oxford: Oxford University Press, 1994.
- Curley, Edwin M. *Descartes Against the Skeptics*. Cambridge, MA: Harvard University Press, 1978.
- Curley, Edwin M. "Descartes, Spinoza and the Ethics of Belief." In Eugene Freeman, ed., *Spinoza: Essays in Interpretation*. La Salle: Open Court Publishing Company, 1975.
- Danto, Arthur, *What Philosophy Is*. New York: Harper & Row, 1968.
- De George, Richard T., ed. *Classical and Contemporary Metaphysics*. New York: Holt, Rinehart and Winston, 1962.
- Dennett, Daniel C. "Descartes's Argument from Design." *Journal of Philosophy* 105:7 (July 2008): 333–45.
- Descartes, René, *Descartes' Conversation with Burman*. Trans. John Cottingham. Oxford: Oxford University Press, 1976.
- Descartes, René. *Descartes: Philosophical Writings*. Trans. and ed. Elizabeth Anscombe and Peter Geach. New York: Macmillan, 1971.
- Descartes, René. *Meditations on First Philosophy with Selections from the Objections and Replies*. Trans. John Cottingham. Cambridge: Cambridge University Press, 1986.
- Descartes, René. *Oeuvres de Descartes*, 11 vols., Rev. ed. Charles Adam and Paul Tannery. Paris: Librairie Philosophique J. Vrin and Le Centre National de la Recherche Scientifique, 1964–1976.
- Descartes, René. *The Philosophical Writings of Descartes*. Trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch, Volumes 1 and 2. Cambridge: Cambridge University Press, 1984–1985.
- Descartes, René. *The Philosophical Writings of Descartes*. Trans. John Cottingham, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny. Volume 3: *The Correspondence*. Cambridge: Cambridge University Press, 1991.
- Descartes, René. *Selected Philosophical Writings*. Trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch. Cambridge: Cambridge University Press, 1988.
- Dicker, Georges. *Berkeley's Idealism: A Critical Examination*. New York: Oxford University Press, 2011.
- Dicker, Georges. *Hume's Epistemology and Metaphysics: An Introduction*. London: Routledge, 1998.

- Dicker, Georges. "Is There A Problem About Perception and Knowledge?" *American Philosophical Quarterly* 15:3 (July 1978): 165–76.
- Dicker, Georges. "Kant's Refutation of Idealism." *Noûs* 47:1 (March 2008): 80–108.
- Dicker, Georges. "Kant's Refutation of Idealism: a Reply to Chignell." *Philosophical Quarterly* 61:242 (January 2011): 175–83.
- Dicker, Georges. "Kant's Refutation of Idealism: Once More unto the Breach." *Kantian Review* 17 (2) (July 2012): 191–95.
- Dicker, Georges. *Kant's Theory of Knowledge: An Analytical Introduction*. New York: Oxford University Press, 2004.
- Dicker, Georges. *Perceptual Knowledge: An Analytical and Historical Study*. Boston: Reidel, 1980.
- Dicker, Georges. "Primary and Secondary Qualities: A Proposed Modification of the Lockean Account." *The Southern Journal of Philosophy*, 15 (Winter 1977): 457–71. Reprinted in Walter E. Creery, ed., *George Berkeley: Critical Assessments*, Volume II, 3 vols. London: Routledge, 1991.
- Dicker, Georges. "Three Questions about *Treatise* 1.4.2." *Hume Studies* 33: 1 (April 2007): 115–53.
- Doney, Willis, ed. *Descartes: A Collection of Critical Essays*. Garden City, NY: Doubleday/Anchor Books, 1967.
- Doney, Willis. "Descartes's Conception of Perfect Knowledge." *Journal of the History of Philosophy* 8 (4) (1970): 387–403. Reprinted in his (ed.) *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Doney, Willis. "The Cartesian Circle." *Journal of the History of Ideas*, 16:3 (June 1955): 324–38. Reprinted in his *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Doney, Willis, ed. *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Ducasse, C.J. "In Defense of Dualism." In Sydney Hook, ed., *Dimensions of Mind*. New York: Macmillan, 1961.
- Feldman, Fred. *A Cartesian Introduction to Philosophy*. New York: McGraw-Hill, 1986.
- Flew, Antony, *An Introduction to Western Philosophy: Ideas and Arguments from Plato to Sartre*. Rev. ed. London: Thames and Hudson, 1989.
- Frankfurt, Harry G. *Demons, Dreamers, and Madmen: The Defense of Reason in Descartes's Meditations*. New York: Bobbs-Merrill, 1970.
- Frankfurt, Harry G. "Descartes' Validation of Reason." *American Philosophical Quarterly* 2:2 (April 1965): 149–56. Reprinted in Willis Doney, ed., *Descartes: A Collection of Critical Essays*. Garden City, NY: Doubleday/Anchor Books, 1967.
- Frankfurt, Harry G. "Memory and the Cartesian Circle." *Philosophical Review* 71 (4) (1962): 504–11. Reprinted in Willis Doney, ed., *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Gewirth, Alan. "The Cartesian Circle." *Philosophical Review* 50 (4) (1941): 368–95. Reprinted in Willis Doney, ed., *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Gewirth, Alan. "The Cartesian Circle Reconsidered." *Journal of Philosophy* 67:19 (1970): 668–85. (1987). Reprinted in Willis Doney, ed., *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Gewirth, Alan. "Descartes: Two Disputed Questions." *Journal of Philosophy* 68:9 (1971), 288–96. Reprinted in Willis Doney, ed., *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Freeman, Eugene, ed. *Spinoza: Essays in Interpretation*. La Salle: Open Court Publishing Company, 1975.

- Grice, Paul H. "The Causal Theory of Perception." Reprinted in Robert J. Swartz, ed., *Perceiving, Sensing, and Knowing*. Berkeley: University of California Press, 1976, pp. 438–72.
- Hatfield, Gary. *Descartes and the Meditations*. London: Routledge, 2003.
- Hick, John. *Evil and the God of Love*. London: Fontana, 1968.
- Hick, John. *Philosophy of Religion*, Second Edition. Englewood Cliffs. NJ: Prentice-Hall, 1973.
- Hintikka, Jaako. "Cogito, Ergo Sum: Inference or Performance." In Willis Doney, ed., *Descartes: A Collection of Critical Essays*. Garden City, NY: Doubleday/Anchor Books, 1967. Also in Alexander Sesonske and Noel Fleming, eds., *Meta-Meditations: Studies in Descartes*. Belmont, CA: Wadsworth Publishing Company, 1965. Originally published in *Philosophical Review* 71 (1) (1962): 3–32.
- Hook, Sydney, ed. *Dimensions of Mind*. New York: Macmillan, 1961.
- Hume, David. *A Treatise of Human Nature*. 2d. ed. Ed. Lewis A. Selby-Bigge and P. H. Nidditch. Oxford: Oxford University Press, 1975.
- Hume, David. *An Enquiry Concerning Human Understanding*. 3d ed. Ed. L.A. Selby-Bigge and P.H. Nidditch. Oxford: Oxford University Press, 1985.
- Hume, David. *Dialogues Concerning Natural Religion*. Indianapolis: Hackett Publishing Company, 1980.
- Kant, Immanuel. *Critique of Pure Reason*. Trans. Norman Kemp Smith. London: Macmillan, 1963.
- Kenny, Anthony. *Descartes: A Study of his Philosophy*. New York: Random House, 1968; reprint ed. New York: Garland, 1987.
- Kenny, Anthony. "Descartes on the Will." In R. J. Butler, ed., *Cartesian Studies*. Oxford: Basil Blackwell, 1972.
- Kenny, Anthony. "The Cartesian Circle and the Eternal Truths." *Journal of Philosophy*, 67 (19) (1970): 685–700.
- Leibniz, Gottfried W. "Critical Remarks Concerning the General Part of Descartes' Principles." In his *Monadology and Other Philosophical Essays*, trans. and ed. Paul Schrecker and Anne Martin Schrecker. New York: Macmillan, 1965.
- Locke, John. *An Essay Concerning Human Understanding*. Abridged and ed. Andrew S. Pringle-Pattison. Oxford: Oxford University Press, 1924.
- Locke, John. *An Essay Concerning Human Understanding*. Ed. Peter H. Nidditch. Oxford: Oxford University Press, 1975.
- Lovejoy, Arthur O. "The Meanings of 'Emergence' and Its Modes." In Richard T. De George, ed., *Classical and Contemporary Metaphysics*. New York: Holt, Rinehart and Winston, 1962. Originally published in *Proceedings of the Sixth International Congress of Philosophy* (1926), 20–33.
- Mackie, John. *The Miracle of Theism*. Oxford: Oxford University Press, 1982.
- Madden, Edward H and Hare, Peter H. *Evil and the Concept of God*. Springfield, IL: Charles C. Thomas, 1968.
- Mandelbaum, Maurice. *Philosophy, Science, and Sense Perception: Historical and Critical Studies*. Baltimore: Johns Hopkins University Press, 1964.
- Matson, Wallace I. *The Existence of God*. Ithaca: Cornell University Press, 1965.
- Montaigne, Michel Eyquem de. *Montaigne's Essays and Selected Writings*. Trans. and ed. Donald M. Frame. New York: St Martin's Press, 1963.
- Naaman-Zauderer, Noa. *Descartes' Deontological Turn*. Cambridge: Cambridge University Press, 2010.
- Nelson, Alan. "Descartes' Ontology of Thought." *Topoi* 16 (2) (1997) 163–78.
- Parkinson, G.H.R., ed. *Routledge History of Philosophy*, Vol. IV. London: Routledge, 1993.

- Price, Henry H. *Thinking and Experience*. Cambridge, MA: Harvard University Press, 1962.
- Prichard, H. A. "Descartes's Meditations." In Willis Doney, ed., *Descartes: A Collection of Critical Essays*. Garden City, NY: Doubleday/Anchor Books, 1967.
- Ragland C. P. "Alternative Possibilities in Descartes's Fourth Meditation." *British Journal for the History of Philosophy* 14:3 (August 2006): 379–400.
- Ragland C. P. "Descartes on the Principle of Alternative Possibilities." *Journal of the History of Philosophy* 44:3 (July 2006): 377–94.
- Rorty, Amélie O., ed. *Essays on Descartes' Meditations*. Berkeley: University of California Press, 1986.
- Rosenthal, David M. "Will and the Theory of Judgment." In Amélie O. Rorty, ed., *Essays on Descartes' Meditations*. Berkeley: University of California Press, 1986.
- Russell, Bertrand. *A History of Western Philosophy*. New York: Simon and Schuster, 1945.
- Russell, Bertrand. *Logic and Knowledge: Essays 1901–1950*. Ed. Robert C. Marsh. New York: Capricorn Books, 1971.
- Schacht, Richard, *Classical Modern Philosophers: Descartes to Kant*. Boston: Routledge & Kegan Paul, 1984.
- Sesonske, Alexander and Noel Fleming, eds., *Meta-Meditations: Studies in Descartes*. Belmont, CA: Wadsworth Publishing Company, 1965.
- Simmons, Alison. "Re-Humanizing Descartes." *Philosophic Exchange* 41 (2010–2011): 52–71.
- Smart, J.J.C. "Sensations and Brain Processes." Reprinted in Clive V. Borst, ed., *The Mind-Brain Identity Theory*. New York: St. Martin's Press, 1970. Originally published in *Philosophical Review* 68 (2) (1959): 141–56.
- Spinoza, Benedictus. *Ethics Demonstrated in Geometrical Order*. In Edwin. M. Curley, ed. and trans., *The Collected Works of Spinoza*, Volume I. Princeton, NJ: Princeton University Press, 1985.
- Swartz, Robert J., ed. *Perceiving, Sensing, and Knowing*. Berkeley: University of California Press, 1965.
- Thomas, Stephen. *Practical Reasoning in Natural Language*. 3d. ed. Englewood Cliffs, NJ: Prentice-Hall, 1986.
- Van Cleve, James. "Foundationalism, Epistemic Principles, and the Cartesian Circle." *Philosophical Review*, 88:1 (January 1979): 55–91. Reprinted in Willis Doney, ed., *Eternal Truths and the Cartesian Circle: A Collection of Studies*. New York: Garland, 1987.
- Van Cleve, James. "On a Little-Noticed Fallacy in Descartes." Manuscript.
- Williams, Bernard. *Descartes: The Project of Pure Enquiry*. Harmondsworth: Penguin Books, 1978.
- Williams, Bernard. "Descartes's Use of Skepticism." In Myles Burnyeat, ed., *The Skeptical Tradition*. Berkeley: University of California Press, 1983.
- Wilson, Margaret. *Descartes*. Boston: Routledge & Kegan Paul, 1978.
- Wolff, Robert Paul. *Kant's Theory of Mental Activity*. Cambridge, MA: Harvard University Press, 1963.
- Wittgenstein, Ludwig. *Philosophical Investigations*, 3d. ed. Trans. Elizabeth Anscombe. Oxford: Blackwell, 1958.

INDEX

- a posteriori* knowledge, 23–24
a priori knowledge, 23–24, 33–34, 37–38,
134
accidents. *See* properties
Alanen, Lilli, 200–202
analytic statements, 33–35, 37–38
animal faith, 279
animal spirits, 306–307
Anscombe, Elizabeth, 109
Anselm, St., 206, 221, 228–229, 231
Aquinas, Thomas, 83, 101, 168–169,
208, 243
Archimedes, 39
argument from change, 50–54, 58–64,
298, 321
argument from doubt, 88–90, 260
argument from ignorance, 260
Aristotle, 5–6, 50, 65, 67, 81–83, 87,
137, 192, 252, 301
Arnauld, Antoine, 144–145, 150, 168,
268–274
astronomy, 5–6
atomism, 283, 285
Augustine, St., 210–211, 215, 217
Ayer, Alfred Jules, 281
basic principles, 14–15
Berkeley, Bishop George, 50, 66, 70, 139,
281, 289–291
Blum, Roland P., 172n49
body, the. *See also* matter; physical world
 clear and distinct perceptions and, 86,
 260
 Cogito and, 75
as distinct from the mind, 86, 90, 258,
260, 263–269, 271–272, 278,
280, 298, 301, 303, 310, 313,
316, 320, 325–326
doubt regarding, 26, 40, 44, 75, 82, 88
existence of, 258, 260, 264–266,
271–272, 280, 302
idea of, 258
interaction with mind and, 274,
301–311, 315, 327
self and, 80, 88–89
soul and, 304, 306–307, 310, 321, 325
Boyle, Sir Robert, 285
Brahe, Tycho, 6
brain
 brain events and, 307–308, 311–313,
 317–318, 321
 mind and, 304–307, 310, 313,
 317–318
 neuroscience and, 38, 307, 311–312,
 317–319
Broad, C.D., 239, 304, 310–311, 317
bundle theory, 50–51
Burman, Frans, 24, 310
Carnap, Rudolph, 244
carnivorous cows example, 237–240, 242
Cartesian circle
 Arnauld on, 144–145, 150, 268
 clear and distinct perceptions and, viii,
 144–145
 creation of eternal truths and,
 166–170, 173
 general rule defense and, 153–164

- Cartesian circle (*continued*)
 memory defense and, viii
 radical doubt of reason and, 164–166,
 169, 171–173
 restriction defense and, 147–154
 validation of reason and, 170–176
- Cartesian dualism. *See also* real
 distinction between mind and
 body
 arguments for, 254
 assessment of, 313–327
 criticisms of, 308–309
 defined, 86
 immortality and, 87, 267
 implications of, 68
 mind and body in, 86, 90, 258, 260,
 263–269, 271–272, 278, 280,
 298, 301–303, 310–311, 313,
 316, 320, 325–326
 physics and, 87
 Platonic thought and, 83
 problem of interaction and, xi, 309–313
- Cartesian Trialism, 302
- Cartwright, Richard L., 237n5
- Caterus, xiii, 243, 248–251, 253
- causal adequacy principle, 117–118, 133,
 135, 143, 146
- causal conception of perception (CCP),
 34–37
- causal principle, 118–119, 122, 139
- cause
 effect and, 115–119, 121–122, 124,
 128–136, 140–143, 310
 efficient and total, 115, 117, 167
 events and, 131–132, 189, 310
 final, 87
 formal and eminent containment and,
 119–122, 278
 of ideas, 104, 110–114, 117–120,
 122–125, 129, 142–144
 intermediate, 311
 performanceist assumption and,
 130–131, 135
 proximate, 311–312
 reality and, 115–126, 128–136,
 140–143
 remote, 311
 of sensory perceptions, 34–36, 38,
 275–282, 286
- centaur example, 230
- certainty. *See also* clear and distinct
 perceptions
 basic principles and, 14
 conditions of observation and, 17
 Descartes's search for, 4, 8–11, 39
 knowledge and, 4
 logic and, 22–23
 mathematics and, 9, 21–23
 negation and, 12–13
 of one's own existence, 40, 259
 of one's own thoughts, 43–44
 senses and, 15–17, 23–24, 31
- Chappell, Vere, 189
- circle and rhombus example, 227–228
- clarity-and-distinctness criterion of
 truth, 114, 126
- clear and distinct perceptions
 assent-compelling nature of, 128,
 143–144, 152–156, 158,
 160–163, 169–170, 186–187,
 190, 192–196, 198–204
- beliefs from, 159–160
 the body and, 86, 260
- Cartesian circle and, viii, 144–145
- Cogito* and, 93–94, 96–97, 263
- compared to moral certainty, 200
- as criterion for truth and, 97–100,
 127–129, 143–151, 153–158,
 161–164, 166–167, 172,
 183–184, 186, 222–224,
 251–253, 255, 261–263, 269, 275
- definition of, 94–95, 97
- degrees of reality and, 139
- doubt and, 99–100, 152–153,
 155–156, 158–165, 173, 176
- drug analogy and, 162–163
- first item of knowledge and, 92–93
- God as guarantor of, 128–129, 144,
 147–148, 150–154, 160–161,
 163, 166, 255, 281
- God's existence and, 97–100, 127,
 129, 150–151, 153, 165, 171,
 175, 222–224, 252–253
- intellect and, 129, 155, 163, 176, 182,
 184, 188, 202, 216
- mathematics and, 95–96, 99, 155,
 157, 159, 222, 269
- principle of alternative possibilities
 and, 195–197, 200
- will and, 186–187, 190, 192–204, 219

- cogitationes* (thoughts), 44, 54, 86, 92
Cogito
additional premise question, 75–79
the body and, 75
clear and distinct perceptions and, 93–94, 96–97, 263
components of, 93, 95–96
criticisms of, 69–71, 74–76
Descartes's belief in the indubitability of, 157–159, 163, 190, 204, 254, 258, 281
“I” premise within, 47–48, 57–58, 69–74, 77
Descartes's various formulations of, 41
dualism and, 75
extraction argument and, 94
performative interpretation of, 48–49
question-begging and, vii, 47–49, 57, 71–72, 74–75, 77
Russell's treatment of, 47
six-step reconstruction of, 54–58
substance theory and, vii, xi, xiii, 49, 53–55, 57–58, 68–70, 79
unreconstructed version of, 71–79, 92, 95
thinking as proof of existence and, 41–43, 71, 73–75, 77–79, 91–97, 99, 113
conceivability, logical possibility and, 264–266, 268, 314, 317
conditions of observation, 17, 19, 29–31, 284–285, 287–290
consciousness, 133, 213–214, 290, 317–319
continuous creation, 138, 169, 324
contradictories, 167, 207, 240–241, 262
contraposition, 13
Copernicus, Nicholas, 6
Copleston, Frederick, 95, 222–223
core argument for God's existence. *See under cosmological argument for God's existence (Descartes)*
Cornman, James, 317–318
cosmological argument for God's existence (Aquinas), 101
cosmological argument for God's existence (Descartes)
core of, 110–114, 122, 124, 128–129
criticisms of, 129–144
degrees of reality and, 125–126
sequel to, 126–129
subargument for first premise of core of, 114–124, 128–129
Cottingham, John
scholarship of, 41, 111–112, 117, 120, 132–133, 135, 143–144, 160–161, 186, 302–303, 307–308
translations of Descartes by, viii, 41, 99n3, 120, 171, 191n10, 249, 259–260
criterion-not-needed strategy, 147, 153–164, 175–176. *See also general rule defense*
Curley, Edwin M., 99, 108n9, 114n12, 142, 184, 259
Darwin, Charles, 6
deceiver argument. *See also God, deception and*
a priori premises of, 34, 37–38
boundary-setting aspect of, 83
causal conception of perception (CCP) and, 34–37
Descartes's response to, 24–25, 40–41, 44, 126–127
“evil demon” and, 26–27, 36–38, 40, 44–45, 80–81
thinking and, 82–83
deception
God and, 24–25, 99–100, 126–129, 143–147, 149–152, 156, 161–163, 165, 170–176, 181–182, 279–281
senses and, 7–8, 15–16, 19–20, 22, 27–32, 38, 46, 103, 128, 163
deflationism, 238–240, 242
degrees of reality
existence of God and, 108, 125–126
finite substances and, 107–108, 137–138, 177
ideas and, 107–114, 129, 142–143, 176–179
infinite substances and, 107, 137–138, 177–178
modes and, 137–138
substance and, 138–139
Democritus, 283
Dennett, Daniel, 180
descriptive predicates, 232–234, 240, 242

- determinable properties, 51–52, 62–64, 299–300
- determinate properties, 51–54, 59, 61–64, 299–300
- Dewey, John, 3, 282
- dispositions, 284–291, 293
- double negation, 13–14, 37
- doubt
- grounds or reasons for, 31–32, 171–175
 - of logic, 22–23
 - of mathematics, 22–24
 - method of, 3, 26, 265–266
 - of one's own body's existence, 26, 40, 44, 75, 82, 88
 - of one's own existence, 40, 75, 80–81, 88
 - of the physical world, 22, 25–26, 39–40, 44, 53, 75, 82, 87–88, 98, 101, 134, 169, 274
 - of the reliability of human reason, 164–166, 169, 171–173
 - of the senses, 31–33, 98, 152, 166, 257
- doxastic attitudes, 11–12
- doxastic voluntarism, 185
- dragon example, 242
- dream argument, 17, 19–22, 24, 27, 30–32, 257
- dreams, 19–22, 30–32, 42, 45, 104
- dualism. *See* Cartesian dualism
- dualistic interactionism, 305, 310, 313–314, 318
- Ducasse, Curt J., 304, 310–311, 313
- Elizabeth, Princess of Bohemia, 309
- eminent reality, 120. *See also* formal and *eminent containment*
- emotions, 190, 192, 308
- epiphenomenalism, 313–314, 317
- epistemology, 24, 38, 79, 264, 266, 322
- error
- deciding to believe and, 184–185
 - evil and, 205
 - fullness of creation and, 215–216
 - intellect and, 182, 219
 - natural impulses and, 104
 - will and, 25, 104, 129, 181–185, 199, 205, 214–217, 219
- essence
- clear and distinct perceptions of, 223–224, 253–254
 - definition and, 252–253
 - extension as, 299–301
 - mathematics and, 229
 - ontological argument for God's existence and, 229–230
 - thinking as, 258–261, 264, 269, 299–300, 320
- eternal truths, creation of, 166–170, 173, 207–208
- events
- brain, 307–308, 311–313, 317–318, 321
 - cause-and-effect relationships and, 131–132, 189, 310
 - mental, 307–308, 311–313, 317–319, 321
 - physical, 309, 311
- evil
- aesthetic conception of, 215–219
 - the afterlife and, 212–213
 - fact/value distinction and, 217
 - fullness of creation and, 217–218
 - God and, 181, 205–213
 - moral evil, 209–210, 214
 - nonmoral evil, 209–214
 - private view of, 214–215, 217–219
 - problem of, 181, 205–212, 218
 - theodicy and, 208–217
 - will and, 209–210
- evil demon. *See under* deceiver argument
- existence
- the body and, 258, 260, 264–266, 271–272, 280, 302
 - compared to subsistence, 238–239, 241
 - doubting one's own, 40, 75, 80–81, 88
 - of God, 25, 118, 126–128, 134, 143–147, 149–153, 161–163, 166–167, 173–176, 179–180, 221, 225–228, 231, 233–236, 242–244, 246–255, 267–268, 280
 - Kant's view of, 231–236, 240
 - necessary existence, 248–250, 252–253
 - of the physical world, xi, 22, 25–26, 39–40, 44, 53, 75, 82, 87–88, 97–98, 101, 134, 169, 254–255, 257, 274, 277–282, 287
 - possible existence, 248

- as a property, 231–237, 242
 real predicate question and, 231–233
 thinking as proof of, 41–43, 47–49,
 71, 73–75, 77–79, 82–83, 88,
 91–97, 99, 106, 113, 258–259,
 263–264
 existential import, 246–247
 extension
 characteristics of, 297
 dreams and, 20
 as quality of physical existence, 40,
 52, 63–64, 85, 102n6, 121,
 263, 270–271, 273, 275–276,
 282–284, 286, 292–301, 303
 wax example and, 85–86
 extraction argument, 94
- Feldman, Fred, 45, 308–309
 final causes, 87
 Flew, Antony, 243–245, 252–253
 formal and eminent containment, 111,
 117, 119–121, 278
 formal mode of speech, 122, 243–248,
 251–253
 formal reality, 109–114, 117–126, 129,
 143, 176–177, 179–180, 277–278
 Frankfurt, Harry G., xiii, 4n1, 24,
 147n33, 163–164, 170–172, 187
 free will. *See* will
 Frege, Gottlob, 239
- Galilei, Galileo, 6–7, 87, 285, 291
 Gassendi, Pierre, 42, 73, 116, 177–180,
 198, 216, 309
 Gaunilo, 229–230
 Geach, Peter, 109
 general conservation principles, 133–135
 general rule defense, 153–164
 geometry. *See* mathematics; triangle
 examples
 Geulincx, Arnold, 309
 Gewirth, Alan, xiii, 170–172
 God. *See also* core argument for God's
 existence (Descartes); ontological
 argument for God's existence
 Anselm's definition of, 206, 221, 228
 deception and, 24–25, 99–100,
 126–129, 143–147, 149–152,
 156, 161–163, 165, 170–176,
 181–182, 279–281
- Descartes's definition of, 221, 228
 eternal truths and, 166–169
 evil and, 181, 205–213
 existence of, 25, 118, 126–128, 134,
 143–147, 149–153, 161–163,
 166–167, 173–176, 179–180,
 221, 225–228, 231, 233–236,
 242–244, 246–255, 267–268, 280
 goodness of, 24–25, 175, 179–180,
 205–214, 216, 218, 231, 248, 272
 as guarantor of clear and distinct
 perceptions, 128–129, 144,
 147–148, 150–154, 160–161,
 163, 166, 255, 281
 idea of, 101, 105–108, 110, 113–114,
 116, 124–125, 129, 143,
 176–177, 180, 221, 226, 248, 251
 infinite nature of, 118, 125, 177–179,
 249
 Judeo-Christianity and, 205–206,
 217–218, 221, 228
 omnipotence of, 21–22, 25, 99–100,
 125, 144, 152, 154, 156, 165,
 168–170, 172–175, 179–180,
 205–209, 212–214, 216, 221,
 231, 233, 247, 249, 262, 272,
 324–325
 omniscience of, 205, 214, 231, 247
 perfection of, 98, 113–114, 116,
 126–127, 129, 143, 174–175,
 180, 211, 214, 221, 224–228,
 231, 246–253, 257, 279–281
 gold statute example, 60–62
 Goldbach's conjecture, 316–317
 Grice, Paul, 35–36
- hallucination, 22, 25–26, 38, 45, 279
 Hatfield, Gary, 102n6, 147–148
 heirloom view of causation, 132, 135
 Hick, John, 209, 211–213
 hierarchy
 dependence and, 138–139
 Descartes's three levels of, 109–110,
 137–139
 medieval philosophy and, 5–6
 Hintikka, Jaakko, 48, 76n34, 265
 Hobbes, Thomas, 57, 66, 137
 Hume, David, 50, 66, 70, 101, 118–119,
 131–132, 140, 142, 164–165,
 206, 281–282, 313

- Hyperaspistes, 70–71, 116
 hypothetical syllogism, 13
- “I”
 Descartes’s use of in the *Meditations*, 14
 special semantics of, 71–72
 reference to the mind and, 320
- I think, therefore I am. *See Cogito*
- ideas
 causes of, 104, 110–114, 117–120, 122–125, 129, 142–144
 degrees of reality and, 107–114, 129, 142–143, 176–179
 formal reality and, 110–114, 122–124, 129, 176–177, 179–180, 277–278
 of God, 101, 105–108, 110, 113–114, 116, 124–125, 129, 143, 176–177, 180, 221, 226, 248, 251
 judgments and, 102–103
 material falsity and, 102n6
 material objects and, 104–106, 112–113, 120, 123, 125
 modes of thought and, 105–106, 112–113, 125
 natural light and, 104
 nature of, 101–103, 105
 objective reality and, 108–114, 118, 122–124, 129, 176–177, 180, 277–278
 perfection and, 107
 pictures and, 101, 111
 representational content of, 123
 self and, 124–125
 truth/falsity and, 102
 types of, 103
- identity through time, 61–63, 300–301
- imagination
 physical world and, 84–85, 257
 self and, 91
 senses and, 84–85
- immortality, 87, 267, 297–298, 319, 321, 325–326
- indifference, Descartes’s conception of, 197–198
- indubitability
 of certain propositions, 9–10, 45, 57, 80, 93, 142, 162–163, 170
 compared to truth, 163–164
- infinite regress, 17, 66, 122
- infinite, human conception of the, 177–179
- inflationism, 238–241
- insanity, viii, 17–18, 46–47
- intellect
 clear and distinct perceptions and, 129, 155, 163, 176, 182, 184, 188, 202, 216
 error and, 182, 219
 will and, 183, 186, 190–191, 193, 202, 214, 219
- Irenaeus, 210–213, 217
- Iseminger, Gary, vii, 72n12
- Kant, Immanuel, 23, 101, 134, 203n23, 231–237, 239–242, 248, 282
- Kenny, Anthony, xii, 57–58, 70–71, 137, 153–154, 156–158, 161, 170n46, 187, 197n15, 270
- Kepler, Johannes, 6, 87
- knowledge. *See also* clear and distinct perceptions
 certainty and, 4
 Descartes’s notions of, 8
 natural light and, 73, 104, 115, 127, 138, 146, 182, 214, 249–250
a posteriori, 23–24
a priori, 23–24, 33–34, 37–38, 134
- Lehrer, Keith, 11
- Leibniz, Gottfried von, 50, 168, 280–281, 309
- Locke, John, 61–62, 101, 281, 283–288, 290–291, 293–295
- Lovejoy, Arthur O., 130, 135
- Mack, Chelsie, 139n25
- Mackie, John L., 133–135
- Malebranche, Nicolas, 309
- Mandelbaum, Maurice, 28
- manifest aspect. *See under* secondary qualities
- masked man fallacy, 89
- material mode of speech, 243–248, 251–252
- material world. *See* physical world
- materialism, 313–314, 316, 318–319
- mathematics
 certainty and, 9, 21–23
 clear and distinct perceptions and, 95–96, 99, 155, 157, 159, 222, 269

- Descartes's knowledge of, 8, 10
doubt of, 22–24
eternal truths and, 166–168
necessary truths and, 96, 99
physics and, 21
a priori knowledge and, 23–24
matter. *See also* physical world
atomic theory of, 286
characteristics of, 86
compared to space, 283, 292–295
mind as distinct from, 53, 68, 87–88,
 97–98, 254–255, 268, 321
solidity and, 294
medieval philosophy, 5–6, 50, 55–56, 69,
 116–117, 130. *See also* Anselm St.;
 Aquinas, Thomas; Augustine, St.
Meinong, Alexius, 238
memory, 29, 147, 149
memory defense, viii, 147
mental events, 307–308, 311–313,
 317–319, 321
Mesland, Denis, 197–199, 202, 204
Mill, John Stuart, 281
mind
 brain and, 304–307, 310, 313,
 317–318
 as distinct from matter, 53, 68, 87–88,
 97–98, 254–255, 268, 321
 as distinct from the body, 86, 90, 258,
 260, 263–269, 271–272, 278,
 280, 298, 301, 303, 310, 313,
 316, 320, 325–326
interaction with body and, 274,
 301–311, 315, 327
mind-body problem, 274, 309. *See also*
 problem of interaction
Moderate Realism, 65, 67–68
Modus Ponens, 23, 29, 37, 74, 76, 236
Modus Tollens, 23, 37, 231–232, 236
Montaigne, Michel, 7–8, 22
Moore's Paradox, 157
mountain and valley example, 225–226

natural light
 clear and distinct perceptions and, 146
 contrasted with nature, 104, 279–280
 doctrine of irresistibility regarding,
 186
existence and, 42
ideas and, 104
knowledge and, 73, 104, 115, 127,
 138, 146, 164, 182, 214, 249–250
reason and, 58
negative existentials, 237–240, 242
Nelson, Alan, 200–203
Neo-Platonists, 107
nerves, 286, 306–308
Newton, Isaac, 6–7
nominalism, 66–68
noncontradiction, principle of, 99, 156,
 158–159, 163, 204
nothingness, 55–56, 116–117, 119,
 135–136, 139–141

objective reality. *See under* ideas
occasionalism, 309–310
Ockham's Razor, 318
omnipotence
 God and, 21–22, 25, 99–100, 125, 144,
 152, 154, 156, 165, 168–170,
 172–175, 179–180, 205–209,
 212–214, 216, 221, 231, 233,
 247, 249, 262, 272, 324–325
logical possibility and, 168–169,
 207–208, 262
ontological argument for God's existence
Caterus's objection to, 248–251, 253
criticisms of, 228–254
Descartes's version of, viii, xii, 25, 101,
 221–228, 254–255
formal mode of speech and, 122,
 243–248, 251–253
Gaunilo's objections to, 229–230
Kant's objection to, viii, xiii, 231–237,
 240–242
material mode of speech and, 243–248,
 251–252
necessary existence and, 248–251
real predicate question, 231–233
ontology, 50, 66–68

perception. *See* clear and distinct
 perceptions; senses
perfect island example, 229–230
philosophy of mind, 314
physical world
 Descartes's view of the nature of the,
 282–301
different conceptions of substance
 and, 295–301

- physical world (*continued*)

dispositions and, 284–291, 293

doubting the existence of, 22, 25–26, 39–40, 44, 53, 75, 82, 87–88, 98, 101, 134, 169, 274

God and, 277–279

imagination and, 84–85, 257

matter and, 277–279

one-substance view of, 64, 297–299, 301

primary qualities and, 283–285, 287–289

proving the existence of, xi, 97–98, 254–255, 257, 274, 277–282, 287, 315

secondary qualities and, 102n6, 283–291, 293

senses and, 276–283, 286–287

solidity and, 283–285, 292–295, 301
- physics

Aristotle and, 5–6, 87

immortality and, 326

mathematics and, 21

Newton and, 6–7

teleology and, 6–7, 87

pineal gland, 306–308

Plato, 65–67, 83, 107, 116, 192, 211, 238

Platonic Realism, 65–68

Plotinus, 211

possible worlds, 96, 168, 194, 249, 252

precontainment principle, 115–116, 130–136, 140–142

preestablished harmony theory, 309

Prichard, H. A., 165

primary qualities, 280, 283–285, 287–289

principle of alternative possibilities, 187, 195–197, 200

Pringle-Pattison, Andrew S., 105–106

private language, 38n14

problem of interaction, 309–313

properties

accidental, 54, 106, 297–298, 300, 309

attributes and, 55

common, 65–66

compared to substance, 68

determinable, 51–52, 62–64, 299–300

determinate, 51–54, 59, 61–64, 299–300

dispositional, 284–285, 288
- essential, 54

nothingness and, 55–56

problems of universals and, 65–68

spatial location and, 66

thoughts as, 56, 64, 68–69, 79
- Ptolemy, 5

Pyrrho, 7, 22
- qualities. *See* primary qualities; properties; secondary qualities

question-begging arguments, 47–49, 57, 71–72, 74–77, 92

Quine, W.V., 33n11
- Ragland, C. P., 187, 201

real distinction between mind and body

Arnauld's exchange with Descartes on, xii, 268–274

Descartes's proof of, 257–258, 260–269, 271–274, 277, 281, 314–316, 319–322, 325

real predicates, 231–233

reason. *See also* clear and distinct perception; natural light

doubt of, 164–166, 169, 171–173

validation of, 170–176

representational theory of perception, 103–104

representative reality. *See* objective reality

res cogitans, 86–87, 303, 322

res extensa, 86–87, 297–298, 303, 322

restriction defense, 147–154

resurrection, 267

Russell, Bertrand, 47–48, 50, 57, 67, 238–239, 281
- Santayana, George, 279

Schacht, Richard, 69–70

science, the rise of, 5–7, 10, 83, 87

scientific realists, 60

secondary qualities, 102n6, 280, 283–291, 293

dispositional aspect of, 290–291

manifest aspect of, 290–291

self, the

body and, 80, 88–89

doubting of the, 80–81

ideas and, 124–125

imagination and, 91

- senses and, 91
 soul and, 80–81, 83, 88
 thinking and, 80, 86, 88–93, 96, 106, 121
 self-refuting argument, xii, 27–33, 37–38
 senses
 a posteriori knowledge and, 23
 causal elements of, 34–36, 38, 275–282, 286
 certainty and, 15–17, 23–24, 31
 deception and, 7–8, 15–16, 19–20, 22, 27–32, 38, 46, 103, 128, 163
 doubt of, 31–33, 98, 152, 166, 257
 dreams and, 19–21, 30–31
 imagination and, 84–85
 physical world and, 276–283, 286–287
 skeptics and, 7–8
 thinking and, 276–277
 simplicity, principle of theoretical, 318–319
 skepticism
 arguments for, 7–8, 10, 22, 29
 Descartes's response to, 8, 15–16
 solidity, 283–285, 292–295, 301
 solipsism, 124–125
 Sosa, Ernest, 230n2
 soul
 Aristotle's notion of, 81–83
 body and, 304, 306–307, 310, 321, 325
 Cartesian dualism and, 87
 development of, 212, 217
 existence of, 266–267, 271n5
 finite nature of, 179
 immortality and, 297–298, 325–326
 Judeo-Christian notion of, 83
 the self and, 80–81, 83, 88
 world-soul notion and, 70–71
 space, 292–295
 spatio-temporal continuity, 61–64, 301
 Spinoza, Benedictus de, 50, 301, 324
 subsistence. *See under* existence
 substance theory
 argument from change and, 50–53, 59, 321
 corollary of, 64–68
 criticisms of, 58–61, 300
 definitions of, 49–50, 323–324
 degrees of reality and, 138–139
 extension and, 63–64
 God and, 323–324, 326
 mental *versus* physical substances and, 321–324
 objections to Descartes's version of, 300, 321
 substance-property principle, 55, 138–139
 supremely perfect being. *See God*
 syllogisms, 13, 74, 77–78
 synthetic statements, 33
 Taj Mahal example, 234–235
 tautologies, 252–253
 teleology, 5–7, 87
 thalers example, 232, 235–237
 theodicy. *See also* evil
 definition of, 208–209, 208–213
 Augustine and, 210–211, 215, 217
 Descartes and, 214–217
 Irenaeus and, 210–213, 217
 thinking
 Descartes's conception of, 84
 indubitable quality of, 44–45, 47, 57–58, 80, 93
 as proof of existence, 41–43, 47–49, 71, 73–75, 77–79, 82–83, 88, 91–97, 99, 106, 113, 258–259, 263–264
 the self and, 80, 86, 88–93, 96, 106, 121
 triangle examples
 Arnauld's use of, 268–270
 Descartes's use of, 148–149, 155, 178, 186, 223, 225, 228–230, 254, 269–272
 Pythagorean property and, 270–274
 unicorn example, 240–242, 246
 universals, the problem of, 65–68
 Van Cleve, James, xii, 100, 136, 145n31, 146n32, 153–154, 156, 170n45, 174n45
 vindication-not-needed-strategy, 147–153, 159, 175–176
 wax
 as a model of the entire physical world, 63, 84–86, 263, 282, 299
 substance theory and, 51–53, 59

- will
 clear and distinct perceptions and, 186–187, 190, 192–204, 219
 degrees of freedom and, 189–190, 197–198
 Descartes' conception of, 187–190
 error and, 25, 104, 129, 181–185, 199, 205, 214–217, 219
 evil and, 209–210
 hard determinism and, 189
 indifference and, 197–198
 intellect and, 183, 186, 190–191, 193, 202, 214, 219
 internal compulsion and, 188–189
 judgment and, 193
“morally speaking” versus “absolutely speaking,” 197–198, 200–203
principle of alternative possibilities and, 187
pursuing the good and, 191–193, 198–199
Williams, Bernard, xii, 73n13, 79n16, 112n11, 118, 137, 141, 153–154, 159–161, 179–180, 193–194
Wilson, Margaret, 170n46, 192, 266
winged horse example, 226–227
withholding belief, 10–12, 14, 26, 39
Wittgenstein, Ludwig, 38n14, 151, 282
Zaaman-Zauderer, Noa, 188, 199–203
Zoroastrianism, 206